



**PCT**  
WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales Büro  
INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

|  |   |  |
|--|---|--|
| <b>(51) Internationale Patentklassifikation 6 :</b><br><br><b>G01N 33/00</b>   | <b>A2</b>   | <b>(11) Internationale Veröffentlichungsnummer: WO 98/45700</b><br><br><b>(43) Internationales Veröffentlichungsdatum:</b> 15. Oktober 1998 (15.10.98) |
| <b>(21) Internationales Aktenzeichen:</b> PCT/DE98/01016<br><br><b>(22) Internationales Anmeldedatum:</b> 8. April 1998 (08.04.98)<br><br><b>(30) Prioritätsdaten:</b><br>197 14 558.2      9. April 1997 (09.04.97)      DE<br><br><b>(71)(72) Anmelder und Erfinder:</b> ENGELS, Joachim, W. [DE/DE]; Feldbergstrasse 1, D-61476 Kronberg (DE). WÖRNER, Karlheinz [DE/DE]; Feldbergstrasse 17, D-63303 Dreieich (DE). FAULSTICH, Konrad [DE/DE]; Brunnenweg 3, D-36115 Ehrenberg (DE). BRILL, Hannelore [DE/DE]; Kastanienweg 7 F, D-61462 Königstein (DE).<br><br><b>(74) Anwalt:</b> SCHULER, Peter, Arnulfstrasse 25, D-80335 München (DE).   | <b>(81) Bestimmungsstaaten:</b> US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><br><b>Veröffentlicht</b><br><i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i> |  |
| <b>(54) Title: METHOD FOR THE MASS SPECTROMETRIC SEQUENCING OF BIOPOLYMERS</b><br><br><b>(54) Bezeichnung: VERFAHREN ZUR MASSENSPEKTROMETRISCHEN SEQUENZIERUNG VON BIOPOLYMEREN</b><br><br><b>(57) Abstract</b><br><br><p>The invention relates to a new method for sequencing biopolymers by mass spectrometry. The sequencing of biopolymers is either lengthy or requires exact determination of the mass of the fragments, which is difficult, especially in the case of long polymers. The speed of hydrolysis of phosphodiester, -peptide or -glycoside bonds with exo/endonucleases, -peptidases, -glycosidases or other hydrolytically acting substances is used in the inventive method for the sequence analysis of nucleic acids or other biopolymers. Separation and detection of the fragments produced take place by mass spectrometry by determining the mass and the different peak intensities. The main advantage of the method is that an exact determination of the mass is no longer necessary, and also that low-resolution mass spectrometers can be used. Furthermore, analysis and separation of fragments is extremely rapid, as there is no electrophoresis and the sequence of modified nucleic acids can be determined. A further advantage is that unmarked nucleic acids can be used and no radioactivity is need. The method can be conducted in parallel with a simultaneous sequence analysis of various nucleic acids, which increases sequencing speed. The method can also be used for detecting and determining organisms (fingerprint, footprint), whereby the finger and footprints are more precise than in previous methods: both cleavage fragments are detected after hydrolysis of a phosphodiester bond, whereas hitherto known methods could detect only a marked fragment. Finally, the method can contribute to elucidation of the secondary structure of nucleic acids by mass spectrometry. These principles can also be applied to sequencing or secondary structure determination of other biopolymers, like, for example, peptides and oligosaccharides.</p> |   |  |

### (57) Zusammenfassung

Es wird ein neues Verfahren zur Sequenzierung von Biopolymeren mit Massenspektrometrie beschrieben. Die Sequenzierung von Biopolymeren ist entweder langwierig oder benötigt eine genaue Massenbestimmung von Fragmenten. Dies ist vor allem für lange Polymere schwierig. Die Geschwindigkeit der Hydrolyse von Phosphodiester-, Peptid- oder Glycosidbindungen mit Exo-/Endonukleasen, -peptidasen, -glycosidasen oder anderen hydrolytisch wirkenden Substanzen wird in unserem Verfahren zur Sequenzanalyse von Nukleinsäuren oder anderen Biopolymeren herangezogen. Trennung und Detektion der erzeugten Fragmente erfolgt mit Massenspektrometrie durch Bestimmung der Masse und unterschiedlichen Peakintensitäten. Der primäre Vorteil der Methode liegt darin, daß keine exakte Massenbestimmung mehr notwendig ist und auch Massenspektrometer mit geringer Auflösung verwendet werden können, Analyse und Trennung der Fragmente extrem schnell sind, da sie elektrophoresefrei sind und die Sequenz modifizierter Nukleinsäuren bestimmt werden kann. Weitere Vorteile sind, daß unmarkierte Nukleinsäuren eingesetzt werden können und keine Radioaktivität benötigt wird. Die Methode kann durch gleichzeitige Sequenzanalyse mehrerer Nukleinsäuren parallelisiert werden, wodurch die Sequenziergeschwindigkeit steigt. Die Methode kann weiterhin zur Erkennung und Bestimmung von Organismen herangezogen werden (Fingerprint, Footprint), wobei Finger- und Footprint genauer sind als bei bisherigen Methoden: Es werden nach der Hydrolyse einer Phosphodiesterbindung beide Spaltfragmente detektiert, während herkömmliche Verfahren nur ein markiertes Fragment detektieren können. Schließlich kann die Methode zur Aufklärung der Sekundärstruktur von Nukleinsäuren mit Massenspektrometrie beitragen. Diese Prinzipien lassen sich auch auf die Sequenzierung bzw. Sekundärstrukturbestimmung anderer Biopolymere anwenden, wie z.B. Peptide und Oligosaccharide.

### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

|    |                              |    |                                   |    |   |    |                                |
|----|------------------------------|----|-----------------------------------|----|---|----|--------------------------------|
| AL | Albanien                     | ES | Spanien                           | LS | Lesotho   | SI | Slowenien                      |
| AM | Armenien                     | FI | Finnland                          | LT | Litauen   | SK | Slowakei                       |
| AT | Österreich                   | FR | Frankreich                        | LU | Luxemburg                                       | SN | Senegal                        |
| AU | Australien                   | GA | Gabun                             | LV | Lettland  | SZ | Swasiland                      |
| AZ | Aserbaidshan                 | GB | Vereinigtes Königreich            | MC | Monaco  | TD | Tschad                         |
| BA | Bosnien-Herzegowina          | GE | Georgien                          | MD | Republik Moldau                                 | TG | Togo                           |
| BB | Barbados                     | GH | Ghana                             | MG | Madagaskar                                      | TJ | Tadschikistan                  |
| BE | Belgien                      | GN | Guinea                            | MK | Die ehemalige jugoslawische Republik Mazedonien | TM | Turkmenistan                   |
| BF | Burkina Faso                 | GR | Griechenland                      | ML | Mali  | TR | Türkei                         |
| BG | Bulgarien                    | HU | Ungarn                            | MN | Mongolei  | TT | Trinidad und Tobago            |
| BJ | Benin                        | IE | Irland                            | MR | Mauretanien                                     | UA | Ukraine                        |
| BR | Brasilien                    | IL | Israel                            | MW | Malawi  | UG | Uganda                         |
| BY | Belarus                      | IS | Island                            | MX | Mexiko  | US | Vereinigte Staaten von Amerika |
| CA | Kanada                       | IT | Italien                           | NE | Niger   | UZ | Usbekistan                     |
| CF | Zentralafrikanische Republik | JP | Japan                             | NL | Niederlande                                     | VN | Vietnam                        |
| CG | Kongo                        | KE | Kenia                             | NO | Norwegen  | YU | Jugoslawien                    |
| CH | Schweiz                      | KG | Kirgisistan                       | NZ | Neuseeland                                      | ZW | Zimbabwe                       |
| CI | Côte d'Ivoire                | KP | Demokratische Volksrepublik Korea | PL | Polen   |    |                                |
| CM | Kamerun                      | KR | Republik Korea                    | PT | Portugal  |    |                                |
| CN | China                        | KZ | Kasachstan                        | RO | Rumänien  |    |                                |
| CU | Kuba                         | LC | St. Lucia                         | RU | Russische Föderation                            |    |                                |
| CZ | Tschechische Republik        | LI | Liechtenstein                     | SD | Sudan   |    |                                |
| DE | Deutschland                  | LK | Sri Lanka                         | SE | Schweden  |    |                                |
| DK | Dänemark                     | LR | Liberia                           | SG | Singapur  |    |                                |
| EE | Estland                      |    |                                   |    |   |    |                                |

## Verfahren zur massenspektrometrischen Sequenzierung von Biopolymeren

### Beschreibung

5

10

15

20

25

30

Die zunehmenden Aktivitäten in der Forschung von Nukleinsäuren, besonders von Ribonukleinsäuren sowie von Peptiden und Oligosacchariden in den letzten Jahren, erfordern eine schnelle Standard-Sequenziermethode, die auch Modifikationen detektieren kann. Bisherige Methoden für die Sequenzierung von RNA beruhen auf zweidimensionalen chromatographischen Methoden<sup>1-3</sup>, Maxam-Gilbert-Sequenzierung<sup>4-7</sup> oder Reverse Transkriptase Sanger Sequenzierung<sup>8,9</sup>. Neuere Entwicklungen benutzen Massenspektrometrie zur Primärstrukturbestimmung<sup>10</sup>. Eine Reihe von Arbeiten in den letzten Jahren haben sich mit dem physikalischen Abbau von Oligonukleotiden beschäftigt, wie z. B. Tandem Massenspektrometrie mit Elektrospray Ionisierung, ESI (CID, collision induced dissociation)<sup>11-13</sup>, enzymatische Reaktionen unter Verwendung von Exonukleasen<sup>10,14</sup> oder Oligonukleotid-Aufbau mit Polymerasen<sup>15,16</sup> und physikalische spontane Fragmentierung wie "nozzle skimmer dissociation" (NS) von ESI generierten Nukleinsäure Ionen<sup>17,18</sup> oder spontane Dissoziation der Nukleinsäuren nach Infrarot-Laser Beschuß einer in einer Matrix kristallisierten Probe<sup>18-24</sup>. Diese Methoden versagen aber bei der Sequenzierung von RNA, da die Nukleotide Uridin (U, 306,17) und Cytidin (C, 305,18) fast gleiche Massen besitzen. Deshalb wurde argumentiert, daß die Sequenzierung von RNA durch Exonuklease Abbau (Verdau) und Detektion der erhaltenen Fragmente mit Massenspektrometrie nicht möglich sei<sup>25</sup>.

Wir beschreiben hier eine Methode, mit der man RNA durch Exonukleaseverdau, Trennung und Detektion der erzeugten Fragmente mit Massenspektrometrie, sequenzieren kann. Die Methode kann besonders nützlich sein für die Primärstrukturbestimmung von RNA (oder auch DNA), die länger als 20 Basen ist oder modifizierten Nukleinsäuren, da die Auflösung von Massenspektrometern ein limitierender Faktor für die Sequenzanalyse längerer Nukleinsäurefragmente darstellt. Sie kann auch wertvoll bei der Bestimmung der Sekundärstruktur von Nukleinsäuren sein. Die Nukleinsäurefragmente werden durch unterschiedliche

Massen, der beim Verdau von Exonukleasen (vornehmlich 5'→3'-Exonuklease aus Kalbsmilz und 3'→5'-Exonuklease aus Schlangengift von *crotalus durissus*) austretenden Nukleotide, massenspektrometrisch bestimmt. U und C werden aber durch unterschiedliche Peakintensitäten im Massenspektrum erkannt. Die Unterschiede in den Peakintensitäten werden durch unterschiedliche Geschwindigkeiten bei der Hydrolyse der Phosphodiesterbindungen durch das Enzym hervorgerufen. Dadurch werden Fragmente, die am 5'-Ende ein C enthalten durch 5'→3'-Phosphodiesterase weniger schnell abgebaut und liegen deshalb in viel größeren Konzentrationen vor als z.B. 5'-U enthaltende Fragmente. Die gleiche Beobachtung wird auch für 5'-A-Fragmente gemacht. Adenosin ist jedoch auch durch seine Masse von den anderen Nukleotiden leicht zu unterscheiden. Es ist möglich mehrere Nukleinsäuren gleichzeitig dieser Enzymkinetik zu unterwerfen, um die Sequenziergeschwindigkeit zu erhöhen. Auch der Einsatz von basenspezifischen Exo-/Endonukleasen kann zur Sequenzanalyse und zur schnellen Erkennung und Bestimmung von Organismen, z. B. Viren herangezogen werden, deren RNA oder DNA einem "Fingerprinting" (Verdau von RNA oder DNA mit Exo-/Endonukleasen, die nicht an jedem Nukleotid basenspezifisch schneiden) oder "Footprinting" (Verdau von RNA oder DNA, die mit Nukleinsäure fremden Molekülen wechselwirken und der Hydrolyse mit Exo-/Endonukleasen ausgesetzt werden) unterworfen wird. Die Trennung und Detektion der erzeugten Fragmente durch Massenspektrometrie kann so z.B. zur Prävention von Seuchen oder zur Bestimmung von Organismen und biologischen Waffen eingesetzt werden. Finger- und Footprint sind genauer als bisherige Methoden: Es werden nach der Hydrolyse einer Phosphodiesterbindung beide Spaltfragmente detektiert, während herkömmliche Verfahren nur das markierte Fragment detektieren können. Eine Sekundärstrukturvorhersage von Nukleinsäuren ist schließlich dadurch möglich, daß die Enzyme oft bevorzugt an einzelsträngigen, linearen Bereichen schneiden. Somit können Domänen, Sekundär- und Tertiärstrukturen wie z. B. 'Hairpins' oder 'internal loops' an ihren doppelsträngigen Bereichen erkannt werden.

Diese Prinzipien lassen sich auch auf die Sequenzierung bzw. Sekundärstrukturbestimmung anderer Biopolymere anwenden, wie z.B. Peptide und Oligosaccharide. Die Methode ist sowohl mit MALDI als auch mit DE-MALDI<sup>26,27</sup> durchführbar.

## Beispiele

**Beispiel 1:** Sequenzierung eines 8mers mit 5'→3' Phosphodiesterase und (siehe Fig. 1)

5

Sequenzierung eines 8mers mit mit RNase CL3 (siehe Fig. 2)

10 **Tabelle 1:** Massen und Sequenzen (5'→3'Richtung) der Sequenzierfragmente eines 8mers, die durch enzymatischen Verdau erzeugt wurden

a) mit 5'→3' Phosphodiesterase (aus Kalbsmilz)

15

| Peak | Sequenz  | Masse [Da] | Massendifferenzen [Da]<br>(Peak x-Peak (x + 1)) |
|------|----------|------------|---|
| 1    | CAUGUGAC | 2503,7     |   |
| 2    | AUGUGAC  | 2199,1     | 304,6   |
| 3    | UGUGAC   | 1871,1     | 328,0   |
| 4    | GUGAC    | 1565,9     | 305,2   |
| 5    | UGAC     | 1221,3     | 344,6   |
| 6    | GAC      | 915,3      | 306,0   |
| 7    | AC       | 571,1      | 344,2   |

20

25

b) mit RNase CL3 (aus Hühnerleber)

| Peak | Sequenz | Masse [Da] | Massendifferenzen [Da]<br>(Peak x-Peak (x + 1)) |
|------|---------|------------|---|
| 2    | AUGUGAC | 2199,8     | /   |

30

**Experimentelles:**

Für alle Beispiele der massenspektrometrischen RNA Sequenzierungen gilt, falls nicht anders angegeben: Linear kontinuierliche MALDI-TOF Massenspektrometrie wurde mit einem Fisons VG TOF spec Massenspektrometer (8mer, 9mere RNA und DNA, 16mer, 22mer, 120mer) und DE-MALDI-TOF Messungen mit einem PerSeptive Biosystems Voyager Massenspektrometer (16mer) durchgeführt, die einen UV Stickstofflaser mit einer Emissionsfrequenz von 337nm enthalten. Die Laser Pulsbreite ist 4 ns. Die Spektren wurden im negativ Modus aufgenommen mit Ausnahme des 16mers und des 32 mers. Diese wurden im positiv Modus gemessen. 2,4,6-Trihydroxyacetophenon/Ammoniumcitrat wurde in allen Sequenzierexperimenten als Matrix verwendet. Herstellung der Matrix: Lösung 1 (2,4,6-Trihydroxyacetophenon gesättigt in Ethanol:Wasser, 1:1) und Lösung 2 (0,1 M Ammoniumcitrat in Wasser ~ pH 5,5) werden im Verhältnis 2:1 gemischt. Enzyme wurden von Boehringer Mannheim bezogen: 5'→3' Phosphodiesterase aus Kalbsmilz: Das Enzym greift das Oligonukleotid am 5'-Ende an und hinterlässt 3' Nukleotide. 3'→5' Phosphodiesterase aus *crotalus durissus*: Das Enzym greift das Oligonukleotid am 3'-Ende an und hinterlässt 5' Nukleotide. RNase CL3 aus Hühnerleber: Das Enzym spaltet RNA bevorzugt an Cp/N-Bindungen und produziert Fragmente mit 3' endständigem Cytidinphosphat. Ap/N- und Gp/N-Bindungen werden viel langsamer hydrolysiert, Up/N-Bindungen sehr selten. RNase CL3/Pufferlösung (denaturierend): 2 µl RNase CL3 (0,2U/µl) + 6 µl 8 M Harnstoff in Wasser resultieren in 8 µl 50 mU/µl Enzymlösung. 3'→5'-Phosphodiesterase/Pufferlösung: 2 µl (4 mU/µl) 3'→5'-Phosphodiesterase + 18 µl 0.1 M Ammoniumcitrat, pH 5,5 resultieren in 20 µl 0.2 mU/µl Enzymlösung. Sequenzen der untersuchten RNA-Stücke waren wie folgt:

8mer: 5'-HO-CAUGUGAC-OH-3';  
 9mer (RNA): 5'-HO-GCAUGUGAC-OH-3';  
 9mer (DNA): 5'-HO-GTCACATGC-OH-3';  
 16mer: 5'-HO-GCGUACAUCUCCCCU-OH-3';  
 22mer: 5'-HO-GCUCUUUUCU\*UUUUUCUUUUC-OH-3'; (U\* = <sup>13</sup>C markiertes Uridin an allen fünf Kohlenstoffatomen des Zuckerbausteins);  
 120mer (5s-ribosomale RNA): 5'-pUGCCUGGCGGCCGUAGCGCGGUGGUCCCAC CUGACCCCAUGCCGAACUCAGAAGUGAAACGCCGUAGCGCCGAUGGUAGUG UGGGGUCUCCCAUGCGAGAGUAGGGAACUGCCAGGCAU-OH-3'.

In allen Experimenten wurden Proben von je 1 µl nach einer Inkubationszeit von 1, 3, 6, 10, 20 und 60 Minuten genommen (es sind meist nicht alle Spektren

gezeigt). Die Proben wurden mit der Matrix im Verhältnis 1:1 gemischt, auf die Probenplatte des Spektrometers pipettiert und ca. 20 Minuten an der Luft getrocknet. Die Trocknungs- und Kristallisationszeit kann durch vorsichtiges Anföhen verkürzt werden. Der enzymatische Verdau stoppt, wenn die Proben mit der Matrix vermischt werden.

5

|                                |             |
|--------------------------------|-------------|
| RNA 8mer (0,1 OD)              | 9,0 $\mu$ l |
| 5'→3'-Phosphodiesterase (24mU) | 6,0 $\mu$ l |

10

|          |              |
|----------|--------------|
| $\Sigma$ | 15,0 $\mu$ l |
|----------|--------------|

|                        |      |
|------------------------|------|
| Inkubationstemperatur: | 22°C |
|------------------------|------|

|                  |           |
|------------------|-----------|
| Inkubationszeit: | 6 Minuten |
|------------------|-----------|

15

|                   |             |
|-------------------|-------------|
| RNA 8mer (0,1 OD) | 9,0 $\mu$ l |
|-------------------|-------------|

|  |             |
|--|-------------|
| RNase CL3/Pufferlösung (400 mU; denaturierend) | 8,0 $\mu$ l |
|--|-------------|

|          |              |
|----------|--------------|
| $\Sigma$ | 17,0 $\mu$ l |
|----------|--------------|

20

|                        |      |
|------------------------|------|
| Inkubationstemperatur: | 50°C |
|------------------------|------|

|                  |            |
|------------------|------------|
| Inkubationszeit: | 10 Minuten |
|------------------|------------|

Beispiel 2: Sequenzierung eines 16mers mit 5'→3' Phosphodiesterase (siehe Fig. 3)

25

Sequenzierung eines 16mers mit 3'→5' Phosphodiesterase (siehe Fig. 4)

Tabelle 2: Massen und Sequenzen (5'→3'Richtung) der Sequenzierfragmente eines 16 mers (aufgenommen mit DE-MALDI)

30

a) mit 5'→3' Phosphodiesterase (aus Kalbsmilz)

| 5  | Peak | Sequenz          | Masse<br>berechnet<br>[Da] | Masse<br>gefunden<br>[Da] | Massendifferenzen [Da]<br>(Peak x-Peak (x + 1)) |
|----|------|------------------|----------------------------|---------------------------|---|
|    | 1    | GCGUACAUCUUCCCCU | 4954,0                     | 4954,2                    |   |
|    | 2    | CGUACAUCUUCCCCU  | 4608,8                     | 4609,1                    | 345,1   |
|    | 3    | GUACAUCUUCCCCU   | 4303,6                     | 4304,0                    | 305,1   |
|    | 4    | UACAUCUUCCCCU    | 3958,4                     | 3959,2                    | 344,8   |
| 10 | 5    | ACAUCUUCCCCU     | 3652,2                     | 3652,7                    | 306,5   |
|    | 6    | CAUCUUCCCCU      | 3323,0                     | 3323,6                    | 329,1   |
|    | 7    | AUCUUCCCCU       | 3017,8                     | 3018,6                    | 305,0   |
|    | 8    | UCUUCCCCU        | 2688,6                     | 2689,1                    | 329,5   |
|    | 9    | CUUCCCCU         | 2382,4                     | 2383,0                    | 306,1   |
| 15 | 10   | UUCCCCU          | 2077,2                     | 2077,7                    | 305,3   |
|    | 11   | UCCCCU           | 1771,0                     | 1771,9                    | 305,8   |
|    | 12   | CCCCU            | 1464,8                     | 1465,7                    | 306,2   |
|    | 13   | CCCU             | 1159,6                     | 1160,2                    | 305,5   |
| 20 | 14   | CCU              | 854,4                      | 855,5                     | 304,7   |



b) mit 3'→5' Phosphodiesterase (aus *crotalus durissus*)

| Peak | Sequenz           | Masse<br>berechnet<br>[Da] | Masse<br>gefunden<br>[Da] | Massendifferenzen<br>[Da]<br>(Peak x-Peak (x + 1)) |
|------|-------------------|----------------------------|---------------------------|--|
| 5    | 1 GCGUACAUCUCCCCU | 4954,0                     | 4954,2                    |  |
|      | 2 GCGUACAUCUCCCC  | 4647,8                     | 4648,5                    | 305,7  |
|      | 3 GCGUACAUCUCCC   | 4342,6                     | 4343,7                    | 304,8  |
|      | 4 GCGUACAUCUCC    | 4037,4                     | 4038,6                    | 305,1  |
|      | 5 GCGUACAUCUUC    | 3732,2                     | 3733,4                    | 305,2  |
| 10   | 6 GCGUACAUCUU     | 3427,0                     | 3427,9                    | 305,5  |
|      | 7 GCGUACAUCU      | 3120,8                     | 3122,2                    | 305,7  |
|      | 8 GCGUACAUC       | 2814,6                     | 2815,7                    | 306,5  |
|      | 9 GCGUACAU        | 2509,4                     | 2510,4                    | 305,3  |
|      | 10 GCGUACA        | 2203,2                     | 2204,4                    | 306,0  |
| 15   | 11 GCGUAC         | 1874,0                     | 1874,9                    | 329,5  |

#### Experimentelles:

|    |   |      |              |
|----|---|------|--------------|
| 20 | RNA 16mer (0,1 OD)                            |      | 15,2 $\mu$ l |
|    | 5'→3'-Phosphodiesterase (24mU)                |      | 6,0 $\mu$ l  |
|    | $\Sigma$                                      |      | 21,2 $\mu$ l |
| 25 | Inkubationstemperatur:                        | 22°C |              |
|    | RNA 16mer (0,1 OD)                            |      | 15,2 $\mu$ l |
|    | 3'→5'-Phosphodiesterase/Pufferlösung (0,6 mU) |      | 3,0 $\mu$ l  |
| 30 | $\Sigma$                                      |      | 18,2 $\mu$ l |
|    | Inkubationstemperatur:                        | 40°C |              |

- Beispiel 3: Sequenzierung eines 22mers mit 5'→3' Phosphodiesterase (siehe Fig. 5)
- 5 Sequenzierung eines 22mers mit 3'→5' Phosphodiesterase (siehe Fig. 6)
- Tabelle 3: Massen und Sequenzen (5'→3' Richtung) der Sequenzierfragmente eines 22mers

a) mit 5'→3'Phosphodiesterase (aus Kalbsmilz)

| Peak | Sequenzen                | Masse<br>(berechnet)<br>[Da] | Masse<br>(gefunden)<br>[Da] | Massendifferenzen [Da]<br>(Peak x-Peak (x + 1)) |
|------|--------------------------|------------------------------|-----------------------------|---|
| 1    | GCUCUUUUUCU*UUUUUCUUUUCC | 6712,1                       | 6712,2*                     |   |
| 2    | CUCUUUUUCU*UUUUUCUUUUCC  | 6366,9                       | 6365,9                      | 346,3   |
| 3    | UCUUUUUCU*UUUUUCUUUUCC   | 6061,7                       | 6060,1                      | 305,8   |
| 4    | CUUUUUUCU*UUUUUCUUUUCC   | 5755,5                       | 5753,4*                     | 306,7   |
| 5    | UUUUUCU*UUUUUCUUUUCC     | 5450,3                       | 5447,2                      | 306,3   |
| 6    | UUUCU*UUUUUCUUUUCC       | 5144,1                       | 5139,6                      | 307,6   |
| 7    | UUUCU*UUUUUCUUUUCC       | 4837,9                       | 4831,5                      | 308,1   |
| 8    | UCU*UUUUUCUUUUCC         | 4531,7                       | 4523,8                      | 307,7   |
| 9    | CU*UUUUUCUUUUCC          | 4225,5                       | 4216,6*                     | 307,2   |
| 10   | U*UUUUUCUUUUCC           | 3920,3                       | 3910,3                      | 306,3   |
| 11   | *UUUUUCUUUUCC            | 3609,1                       | 3599,0                      | 310,3   |
| 12   | UUUUUCUUUUCC             | 3302,9                       | 3291,9                      | 307,1   |
| 13   | UUUCUUUUCC               | 2996,7                       | 2984,5                      | 307,4   |
| 14   | UUUCUUUUCC               | 2690,5                       | 2677,3                      | 307,2   |
| 15   | UCUUUUCC                 | 2384,3                       | 2370,9                      | 306,4   |
| 16   | CUUUUUCC                 | 2078,1                       | 2062,8                      | 308,1   |
| 17   | UUUUCC                   | 1772,9                       | 1757,0                      | 305,8   |
| 18   | UUUC                     | 1467,7                       | 1448,6                      | 308,4   |
| 19   | UUC                      | 1161,5                       | 1141,4                      | 307,2   |

b) mit 3'→5' Phosphodiesterase (aus *Crotalus durissus*)

| Peak | Sequenzen                | Masse<br>(berechnet)<br>[Da] | Masse<br>(gefunden)<br>[Da] | Massendifferenzen [Da]<br>(Peak x-Peak (x + 1)) |
|------|--------------------------|------------------------------|-----------------------------|---|
| 1    | GCUCUUUUUCU*UUUUUCUUUUCC | 6712,1                       | 6711,1*                     |   |
| 2    | GCUCUUUUUCU*UUUUUCUUUUUC | 6406,9                       | 6407,3                      | 303,8   |
| 3    | GCUCUUUUUCU*UUUUUCUUUUU  | 6101,7                       | 6101,4                      | 305,9   |
| 4    | GCUCUUUUUCU*UUUUUCUUUU   | 5795,5                       | 5796,8                      | 304,6   |
| 5    | GCUCUUUUUCU*UUUUUCUUUU   | 5489,3                       | 5492,2                      | 304,6   |
| 6    | GCUCUUUUUCU*UUUUUCUUUU   | 5183,1                       | 5184,5                      | 307,7   |
| 7    | GCUCUUUUUCU*UUUUUCUUUU   | 4876,9                       | 4877,8                      | 306,7   |
| 8    | GCUCUUUUUCU*UUUUUCUUUU   | 4571,7                       | 4573,7                      | 304,1   |
| 9    | GCUCUUUUUCU*UUUUUCUUUU   | 4265,5                       | 4266,0                      | 307,7   |
| 10   | GCUCUUUUUCU*UUUUUCUUUU   | 3959,3                       | 3958,9                      | 307,1   |
| 11   | GCUCUUUUUCU*UUUUUCUUUU   | 3653,1                       | 3652,1                      | 306,8   |
| 12   | GCUCUUUUUCU*UUUUUCUUUU   | 3346,9                       | 3345,3                      | 306,8   |
| 13   | GCUCUUUUUCU*UUUUUCUUUU   | 3040,7                       | 3037,9                      | 307,4   |
| 14   | GCUCUUUUUCU*UUUUUCUUUU   | 2729,5                       | 2725,8                      | 312,1   |
| 15   | GCUCUUUUUCU*UUUUUCUUUU   | 2424,3                       | 2420,0                      | 305,8   |
| 16   | GCUCUUUUUCU*UUUUUCUUUU   | 2118,1                       | 2112,1                      | 307,9   |
| 17   | GCUCUUUUUCU*UUUUUCUUUU   | 1811,9                       | 1805,7                      | 306,4   |
| 18   | GCUCUUUUUCU*UUUUUCUUUU   | 1505,7                       | 1498,6                      | 307,1   |
| 19   | GCUCUUUUUCU*UUUUUCUUUU   | 1199,5                       | 1191,7                      | 306,9   |
| 20   | GCUCUUUUUCU*UUUUUCUUUU   | 894,3                        | 886,1                       | 305,6   |
| 21   | GCUCUUUUUCU*UUUUUCUUUU   | 588,1                        | 578,8                       | 307,3   |

**Experimentelles:**

|   |                                |                    |
|---|--------------------------------|--------------------|
| 5 | RNA 22mer (0,1 OD)             | 9,0 $\mu$ l        |
|   | 5'→3'-Phosphodiesterase (24mU) | 6,0 $\mu$ l        |
|   | $\Sigma$                       | <hr/> 15,0 $\mu$ l |

Inkubationstemperatur: 22°C

10

|    |   |                    |
|----|---|--------------------|
| 15 | RNA 22mer (0,1 OD)                            | 9,0 $\mu$ l        |
|    | 3'→5'-Phosphodiesterase/Pufferlösung (0,6 mU) | 3,0 $\mu$ l        |
|    | $\Sigma$                                      | <hr/> 12,0 $\mu$ l |

Inkubationstemperatur: 40°C

20 Beispiel 4: Simultane Sequenzierung von zwei Oligoribonukleotiden (8mer und 22mer; siehe Fig. 7)

Bezüglich Massen und Sequenzen der Sequenzierfragmente siehe Tabellen 1a und 3a.

25

**Experimentelles:**

|    |                                |                    |
|----|--------------------------------|--------------------|
| 30 | RNA 22mer (0,1 OD)             | 9,0 $\mu$ l        |
|    | RNA 8mer(0,1 OD)               | 9,0 $\mu$ l        |
|    | 5'→3'-Phosphodiesterase (20mU) | 5,0 $\mu$ l        |
|    | $\Sigma$                       | <hr/> 23,0 $\mu$ l |

35 Inkubationstemperatur: 22°C  
Inkubationszeit: 20 Minuten

**Beispiel 5:** Fingerprint einer 5s-ribosomalen RNA (120mer) mit RNase CL3  
(siehe Fig. 8)

5 **Experimentelles:**

|  |              |
|--|--------------|
| RNA 120mer (1 OD)                              | 10,0 $\mu$ l |
| RNase CL3/Pufferlösung (400 mU; denaturierend) | 8,0 $\mu$ l  |

|             |                               |
|-------------|-------------------------------|
| 10 $\Sigma$ | <u>18,0 <math>\mu</math>l</u> |
|-------------|-------------------------------|

|                        |            |
|------------------------|------------|
| Inkubationstemperatur: | 50°C       |
| Inkubationszeit:       | 15 Minuten |

15

**Beispiel 6:** Sequenzierung eines 16mer Oligoribonukleotids (Fingerprint) mit  
RNase CL3 (Sequenz aus Tabelle 2;  
siehe Fig. 9)

20

**Experimentelles:**

|  |              |
|--|--------------|
| RNA 16mer (0,1 OD)                             | 9,0 $\mu$ l  |
| RNase CL3/Pufferlösung (100 mU; denaturierend) | 15,2 $\mu$ l |

25

|          |                               |
|----------|-------------------------------|
| $\Sigma$ | <u>24,2 <math>\mu</math>l</u> |
|----------|-------------------------------|

|                        |          |
|------------------------|----------|
| Inkubationstemperatur: | 50°C     |
| Inkubationszeit:       | 1 Minute |

30

Beispiel 7: Simultansequenzierung zweier 9mere (DNA und RNA; siehe Fig. 10)

- 5      Tabelle 4: Fragmente und Massen des Verdaus eines RNA 9mers (5'-HO-GCAUGUGAC-OH-3') mit 3'→5'-Phosphodiesterase (aus *crotalus durissus*)

| 10 | Peak | Sequenz   | Masse<br>(berechnet)<br>[Da] | Masse<br>(gefunden)<br>[Da] | Massendifferenzen<br>[Da]<br>(Peak x-Peak (x + 1)) |
|----|------|-----------|------------------------------|-----------------------------|--|
|    | 1    | GCAUGUGAC | 2854,8                       | 2855,2                      | /  |
|    | 2    | GCAUGUGA  | 2549,6                       | 2549,2                      | 306,0  |
|    | 3    | GCAUGUG   | 2220,4                       | 2220,4                      | 328,8  |
|    | 4    | GCAUGU    | 1875,2                       | 1875,0                      | 345,4  |
| 15 | 5    | GCAUG     | 1569,0                       | 1568,8                      | 306,2  |
|    | 6    | GCAU      | 1223,8                       | 1223,4                      | 345,4  |
|    | 7    | GCA       | 917,6                        | 917,2                       | 306,2  |

- 20      Tabelle 5: Fragmente und Massen des Verdaus eines DNA 9mers (5'-HO-d(GTCACATGC)-OH-3') mit 3'→5'-Phosphodiesterase (aus *crotalus durissus*)

| 25 | Peak | Sequenz      | Masse<br>(berechnet)<br>[Da] | Masse<br>(gefunden)<br>[Da] | Massendifferenzen<br>[Da]<br>(Peak x-Peak (x + 1)) |
|----|------|--------------|------------------------------|-----------------------------|--|
|    | 1    | d(GTCACATGC) | 2698,8                       | 2699,8                      | /  |
|    | 2    | d(GTCACATG)  | 2409,6                       | 2410,1                      | 289,7  |
|    | 3    | d(GTCACAT)   | 2080,4                       | 2080,4                      | 329,6  |
|    | 4    | d(GTCACA)    | 1776,2                       | 1776,6                      | 303,8  |
| 30 | 5    | d(GTCAC)     | 1463,0                       | 1462,9                      | 313,7  |
|    | 6    | d(GTCA)      | 1173,8                       | 1173,4                      | 289,5  |

Tabelle 6: DNA-Abgangsgruppen:

|   |        |        |        |        |
|---|--------|--------|--------|--------|
| 5 | dG     | dA     | dC     | T      |
|   | 329,20 | 313,20 | 289,18 | 304,20 |

10 Tabelle 7: RNA-Abgangsgruppen:

|    |        |        |        |        |
|----|--------|--------|--------|--------|
|    | G      | A      | C      | U      |
| 15 | 345,20 | 329,20 | 305,18 | 306,17 |

**Experimentelles:**

|    |  |                               |
|----|--|-------------------------------|
| 20 | RNA (0,1 OD)                               | 12,5 $\mu$ l                  |
|    | DNA (0,1 OD)                               | 2,3 $\mu$ l                   |
|    | 3'→5'-Phosphodiesterase/Pufferlösung (2mU) | 10,0 $\mu$ l                  |
|    | $\Sigma$                                   | <u>24,8 <math>\mu</math>l</u> |
| 25 | Inkubationstemperatur:                     | 40°C                          |

Im folgenden werden die Vorteile der Methode nochmals aufgeführt:

- 30 - Das Verfahren arbeitet elektrophoresefrei und ist damit sehr schnell.
- Es wird kein zusätzlicher Marker zur Detektion benötigt auch keine Radioaktivität. Damit entfallen alle Markierungsschritte.
- 35 - Das Verfahren nutzt nicht nur die Bestimmung der Masse zur Dateninterpretation, sondern auch die Peakintensitäten. Damit sind auch Spektrometer mit geringer Auflösung, die auch leicht bedienbar sind, einsetzbar und das Verfahren wird kostengünstig.



- Es ist keine genaue Massenbestimmung nötig. Deshalb können auch sehr lange Polymere sequenziert werden.
- 5 - Die Methode kann zur Sekundärstrukturvorhersage von Biopolymeren herangezogen werden.
- Das Verfahren kann zur Sequenzierung von modifizierten Biopolymeren dienen.
- 10 - Das Verfahren kann zur Bestimmung von Organismen dienen (Fingerprint/Footprint).
- Die Methode ist automatisierbar und parallelisierbar.

## Literatur

- 5 (1) Sanger, F., Brownlee, G. G. and Barell, B. G., J. Mol. Biol., (1965) 13, 373
- (2) Brownlee, G. G. and Sanger, F., Eur. J. Biochem., (1969) 11, 395
- (3) Silberklang, M., Gillum, A. M. and RajBhandary, U. L., in Methods in Enzymology. Wu, R. and Grossmann, L. (eds), Academic Press Inc., London and New York, (1979) 59, 58
- 10 (4) Maxam, A. M. and Gilbert, W., Proc. Natl. Acad. Sci. USA, (1977) 74, 560-564
- (5) Stahl, D. A., Krupp, G. and Stackebrandt, E., in Nucleic Acids Sequencing, a practical approach, Howe, C. J. and Ward, E. S. (eds), IRL Press at Oxford University Press, Oxford, New York, Tokyo, (1989) 137
- 15 (6) Waldmann, R., Gross, H. J. and Krupp, G., Nucleic Acids Res., (1987) 15, 7209.
- (7) Zhang, Y., Liu, W. Feng, Y. and Wang, T. P., Anal. Biochem., (1987) 163, 513
- (8) Sanger, F. Nicklen, S., Coulson, A. R., Proc. Natl. Acad. Sci. USA, (1977) 74, 5463-5467
- 20 (9) Hahn, C. S., Strauss, E., G. and Strauss, J., H. in Methods of Enzymology, 180, 121
- (10) Pielles, U., Zürcher, W., Schär, M., Moser, H. E., Nucleic Acids Res., (1993) 21, 3191- 3196
- 25 (11) McLuckey, S. A., Habibi-Goudarzi, S., J. Am. Chem. Soc., (1993) 115, 12085-12095
- (12) Wolter, M. A., Engels, J. W., Eur. Mass Spectrom., (1995) 1, 583-590,
- (13) Ni, J., Pomerantz, S.C., Rozenski, J., Zhang, Y. and McCloskey, J. A., Anal. Chem., (1996) 68, 1989-1999
- 30 (14) Limbach, P. A., McCloskey, J. A., Crain, P. F., Nucleic Acids Res. Symp. Ser., (1994) 31, 127-128
- (15) Roskey, M. T., Juhasz, P., Smirnov, I. P., Takach, E. J., Martin, S. A. and Haff, L. A., Proc. Natl. Acad. Sci. USA, (1996) 93, 4724-4729
- 35 (16) Köster, H., Tang, K., Fu, D.-J., Braun, A., van den Boom, D., Smith, C. L., Cotter, R. J. and Cantor, C.R., Nature Biotechnology, (1996) 14, 1123-1128
- (17) Loo, J. A., Udseth, H., R., Smith, R., D., Rapid Commun. Mass Spectrom. (1988) 2, 207-210
- (18) Little, D. P., Chorush, R. A., Speir, J. P., Senko, M. W., Kelleher, N. L.,

- McLafferty, F. W., J. Am. Chem. Soc. (1994) 116, 4893-4897
- (19) M. Karas, F. Hillenkamp, Anal. Chem., (1988) 60, 2299
- (20) Little, D. P., Speir, J. P., Senko, M. W., O'Connor, P. B., McLafferty, F. W., Anal. Chem., (1994) 66, 2809-2815
- 5 (21) Nordhoff, E., Karas, M., Cramer, R., Hahner, S., Hillenkamp, F., Kirpekar, F., Lezius, A., Muth, J., Meier, C., Engels, J. W., J. Mass Spectrom., (1995) 30, 99-112
- (22) Little, D. P., McLafferty, F. W., J. Am. Chem. Soc., (1995) 117, 6783-6784
- 10 (23) Wu, K. J., Shaler, T. A. and Becker, H., Anal. Chem., (1994) 66, 1637-1645
- (24) Nordhoff, E., Cramer, R., Karas, M., Hillenkamp, F., Kirpekar, F., Kristiansen, K. and Roepstorff, P., Nucleic Acids Res., (1993) 21, 3347-3357
- 15 (25) Kirpekar, F., Nordhoff, E., Kristiansen, K., Roepstorff, P., Lezius, A., Hahner, S., Karas, M. and Hillenkamp, F., Nucleic Acids Research, (1994) 22, 3866
- (26) Wiley, W.C., McLaren, I. H., Rev. Sci. Instrum., (1955) 26, 1150-1157
- (27) Juhasz, P., Roskey, M. T., Smirnov, I. P., Haff, L. A., Vestal, M. L. and  
20 Martin, S. A., Anal. Chem., (1996) 68, 941-946

### Patentansprüche

Wir beanspruchen:

- 5 1. Ein Verfahren zur Sequenzierung von Biopolymeren mit Massenspektrometrie, vornehmlich Ribonukleinsäuren (RNA) oder Nukleinsäuren (DNA) oder Peptiden oder Oligosacchariden durch Verdau des zu untersuchenden RNA- oder DNA- oder Peptid- oder Oligosaccharid- Stranges mit spezifischen Exo-/Endonukleasen, -peptidasen, -carboxyesterasen, -amidasen oder  
10 -glycosidasen oder anderen sequenz- oder basenspezifisch spaltenden Verbindungen, wobei die Trennung und Detektion der erzeugten Fragmente einer Kinetik folgend durch Massenspektrometrie, vornehmlich MALDI, erfolgt und unterschiedliche Peakintensitäten, hervorgerufen durch enzymatische oder chemische Hydrolyse der entsprechenden einzelnen Bindungen, in den  
15 Massenspektren zur Interpretation der Sequenzdaten herangezogen werden.
2. Die Anwendung des Verfahrens nach Anspruch 1 zur Sequenzierung von modifizierter und unmodifizierter DNA, RNA, Peptide und Oligosaccharide. Modifikationen können bei Nukleosiden und Nukleotiden -verglichen mit  
20 Adenosin, Guanosin, Uridin, Cytidin und Thymidin bzw. deren Phosphate und Oligomere- Veränderungen an der Base, am Zucker oder an der Phosphatgruppe bzw. des Phosphatrückgrates betreffen, bei Peptiden Abweichungen der als '20 natürliche Aminosäuren' bekannten Monomere von Ppetiden und bei Oligosacchariden den bekannten 'natürlichen' Sacchariden.
- 25 3. Die Verwendung von Enzymen oder anderen sequenz- oder basenspezifisch spaltenden Verbindungen, die bei der Hydrolyse von Bindungen unterschiedliche Geschwindigkeiten aufweisen und zur Sequenzierung von Nukleinsäuren (RNA/DNA), Peptiden oder Oligosacchariden in einem Verfahren  
30 nach Anspruch 1 und 2 eingesetzt werden.
4. Ein Verfahren zur simultanen Sequenzierung von mehreren Ribonukleinsäuren (RNA), Nukleinsäuren (DNA), Peptiden und Oligosacchariden nach Anspruch 1,  
35 d a d u r c h g e k e n n z e i c h n e t, daß  
der Verdau des zu untersuchenden RNA-, DNA-, Peptid- oder Oligosaccharid-Stranges mit spezifischen Exo-/Endonukleasen, -peptidasen, -amidasen, -carboxyesterasen oder Exo- und Endoglycosidasen oder anderen sequenz- oder basenspezifisch spaltenden Verbindungen gleichzeitig erfolgt, wobei die

Trennung und Detektion der erzeugten Fragmente einer Kinetik folgend durch Massenspektrometrie, vornehmlich MALDI, erfolgt und unterschiedliche Peakintensitäten, hervorgerufen durch enzymatische oder chemische Hydrolyse der entsprechenden einzelnen Bindungen, in den Massenspektren zur Interpretation der Sequenzdaten herangezogen werden.

5

5. Ein Verfahren zur Sequenzierung von RNA, DNA, Peptiden und Oligosacchariden durch Verdau des zu untersuchenden Stranges mit spezifischen Endonukleasen, -peptidasen, -carboxyesterasen, -amidasen, Endoglycosidasen und anderen sequenz- oder basenspezifisch Endospaltenden Verbindungen, wobei die Trennung und Detektion der erzeugten Fragmente durch Massenspektrometrie, vornehmlich MALDI, erfolgt.
6. Ein Verfahren zur Analyse und Bestimmung von Organismen nach Anspruch 1, dadurch gekennzeichnet, daß deren RNA oder DNA einem "Fingerprinting" oder "Footprinting" unterworfen werden und die Trennung und Detektion der erzeugten Fragmente durch Massenspektrometrie, vornehmlich MALDI, erfolgt.
7. Ein Verfahren zur Sekundärstrukturbestimmung von Nukleinsäuren (DNA/RNA), Peptiden und Oligosacchariden durch Verdau der Biopolymeren mit Exo-/Endonukleasen, -peptidasen, -carboxyesterasen, -amidasen und -glycosidasen oder anderen sequenz- oder basenspezifisch spaltenden Verbindungen und Trennung und Detektion der erzeugten Fragmente mit Massenspektrometrie, vornehmlich MALDI, nach Anspruch 1, dadurch gekennzeichnet, daß Sekundärstrukturbereiche (z.B. loops, doppelsträngige Bereiche) nur langsam oder gar nicht bzw. sehr viel schneller von diesen Enzymen angegriffen werden.
8. Kombination der Verfahren nach Anspruch 1-7 zur Analyse von (Ribo-) Nukleinsäuren, Peptiden und Oligosacchariden.

10

15

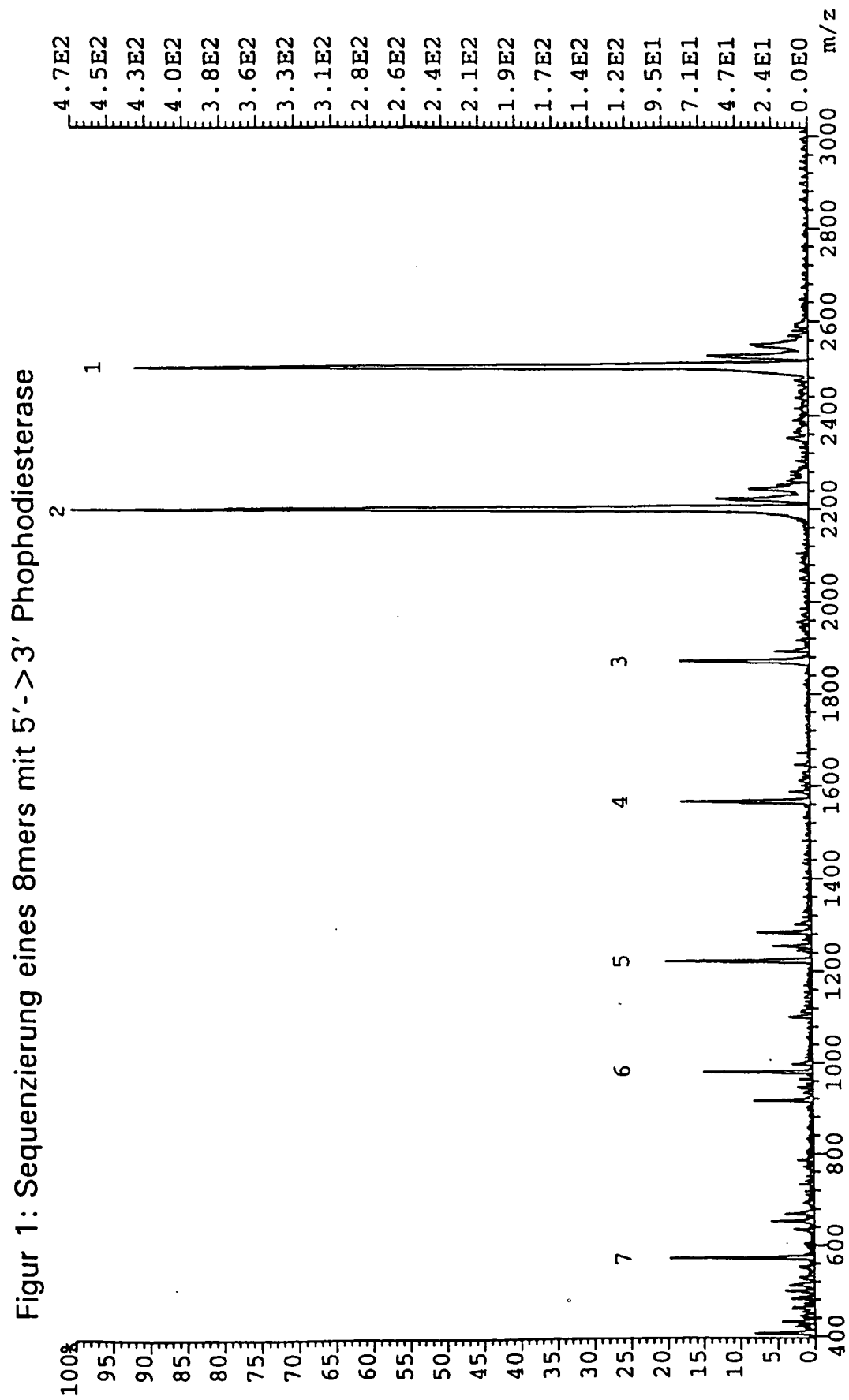
20

25

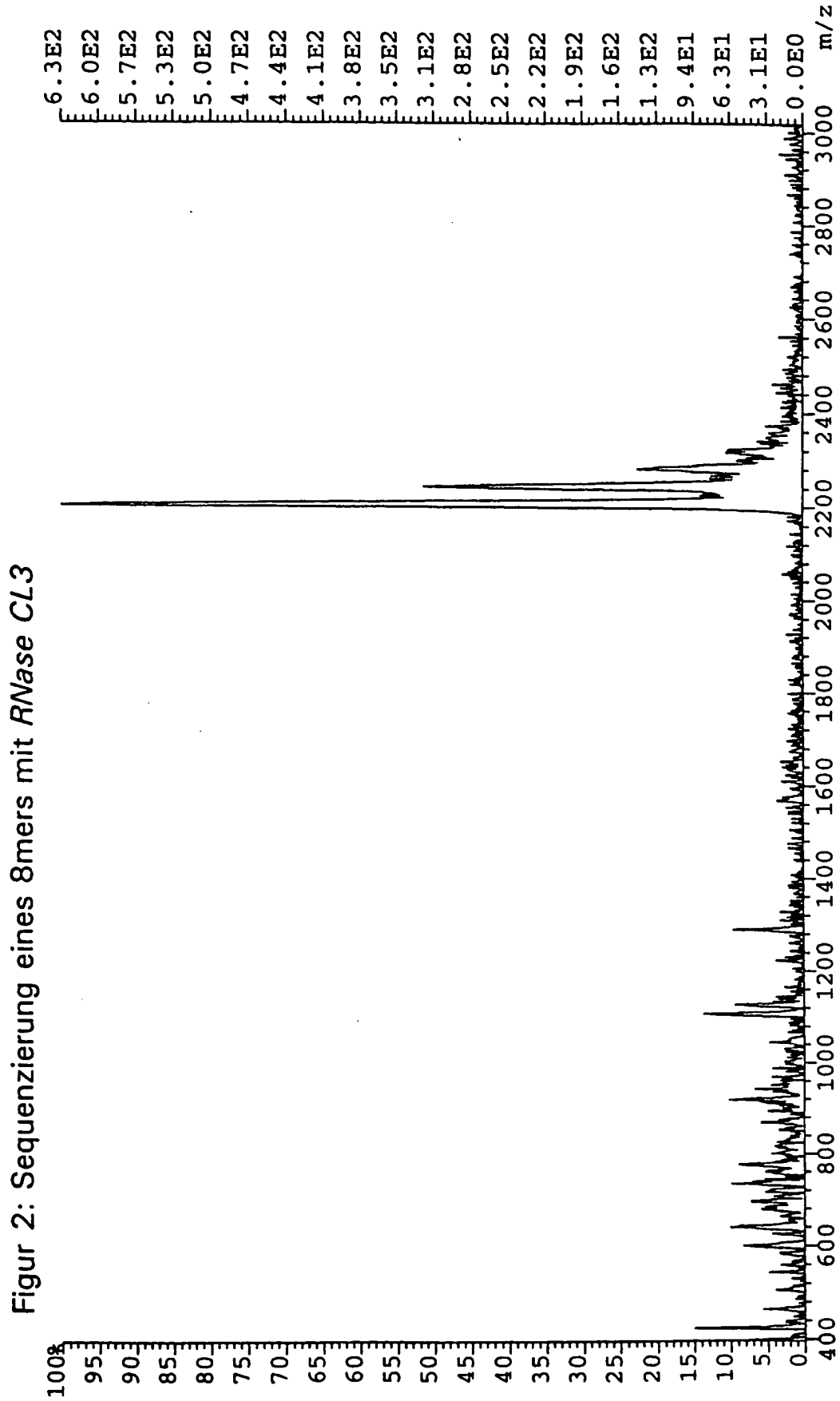
30

35

1/17

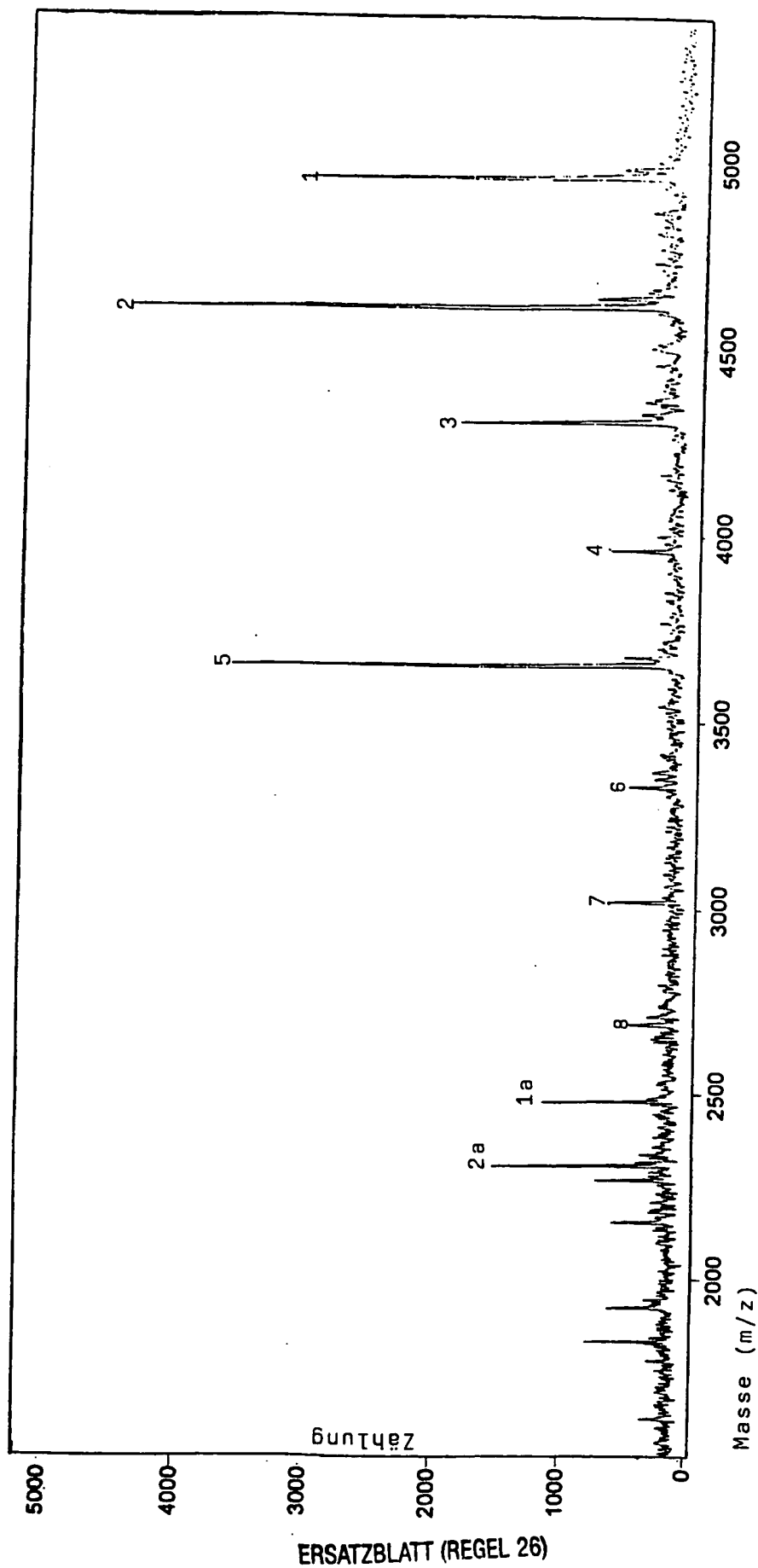


2/17



3/17

Figur 3a: Sequenzierung eines 16mers mit 5' → 3' Phosphodiesterase  
3 Minuten Inkubation

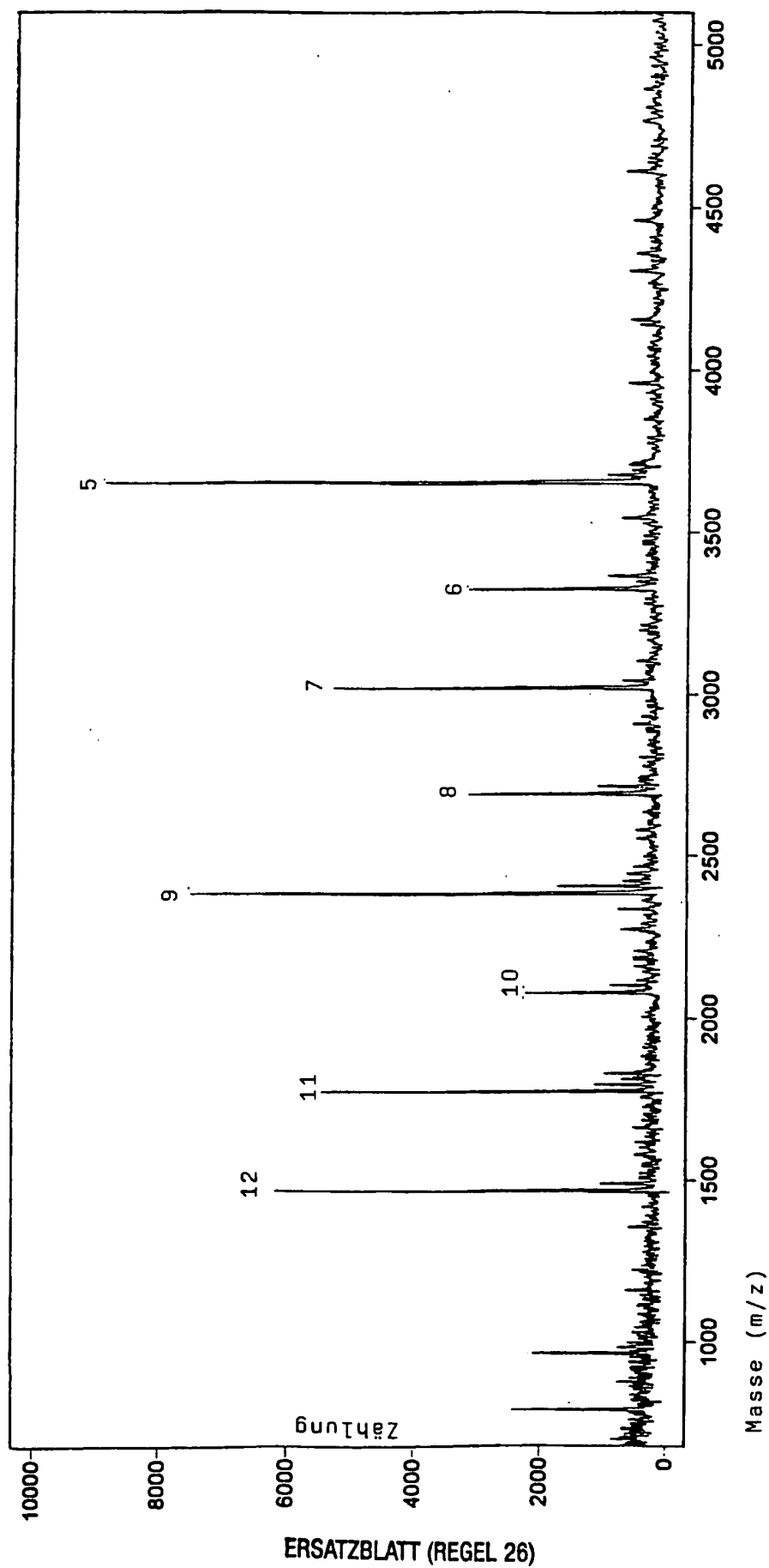


ERSATZBLATT (REGEL 26)



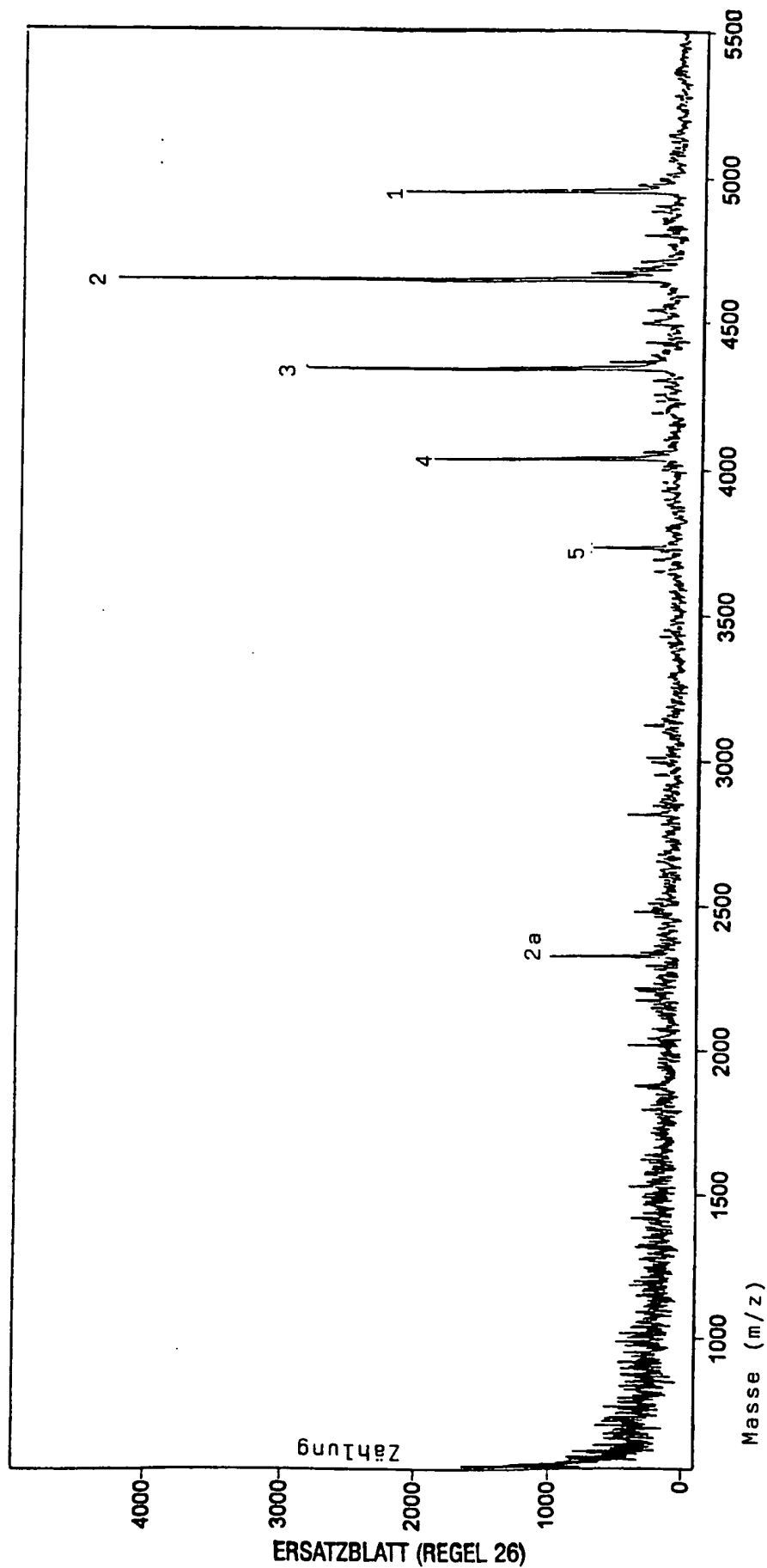
4/17

Figur 3b: Sequenzierung eines 16mers mit 5'->3' Phosphodiesterase  
10 Minuten Inkubation



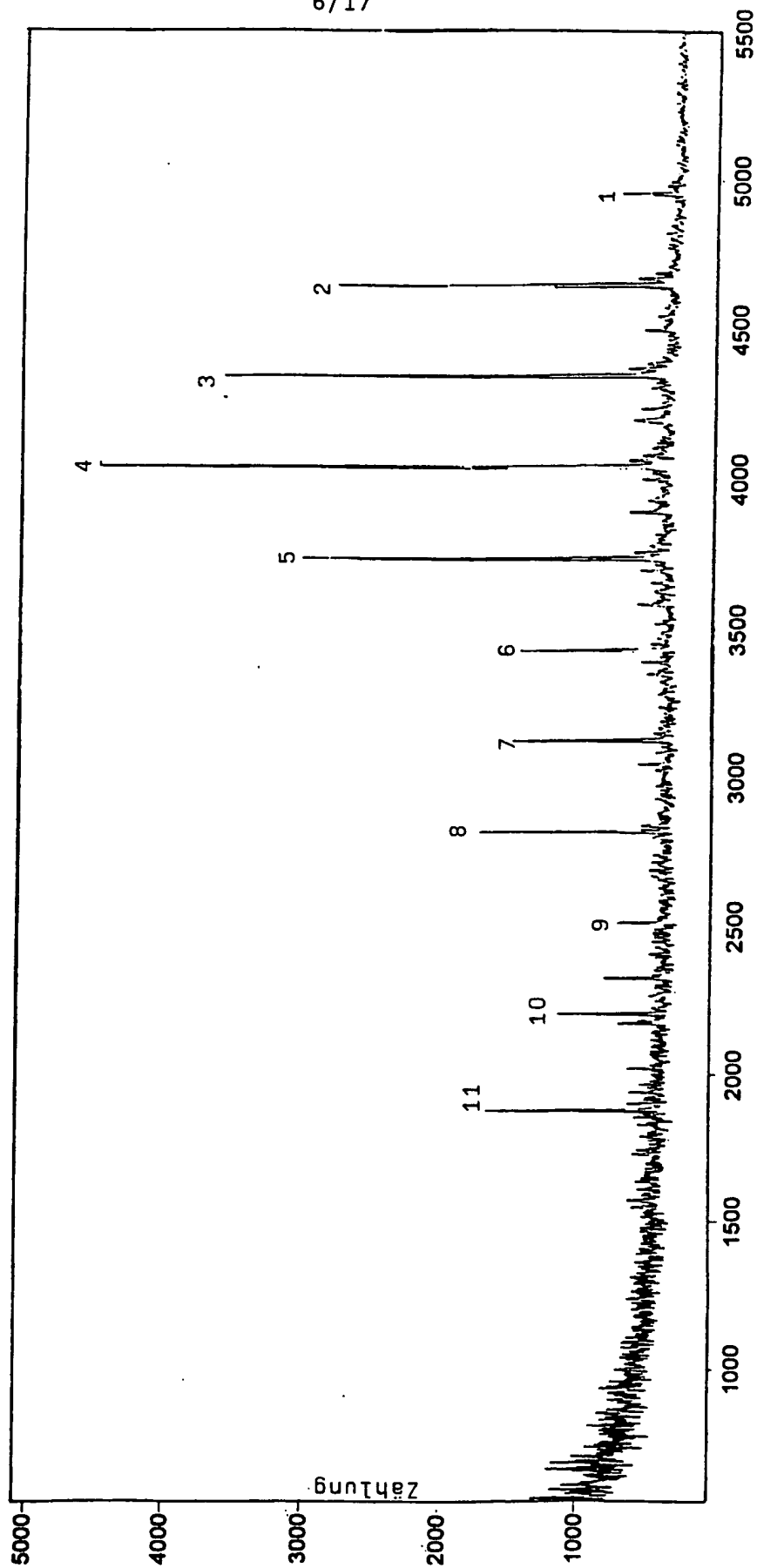
5/17

Figur 4a: Sequenzierung eines 16mers mit 3'->5' Phosphodiesterase  
20 Minuten Inkubation



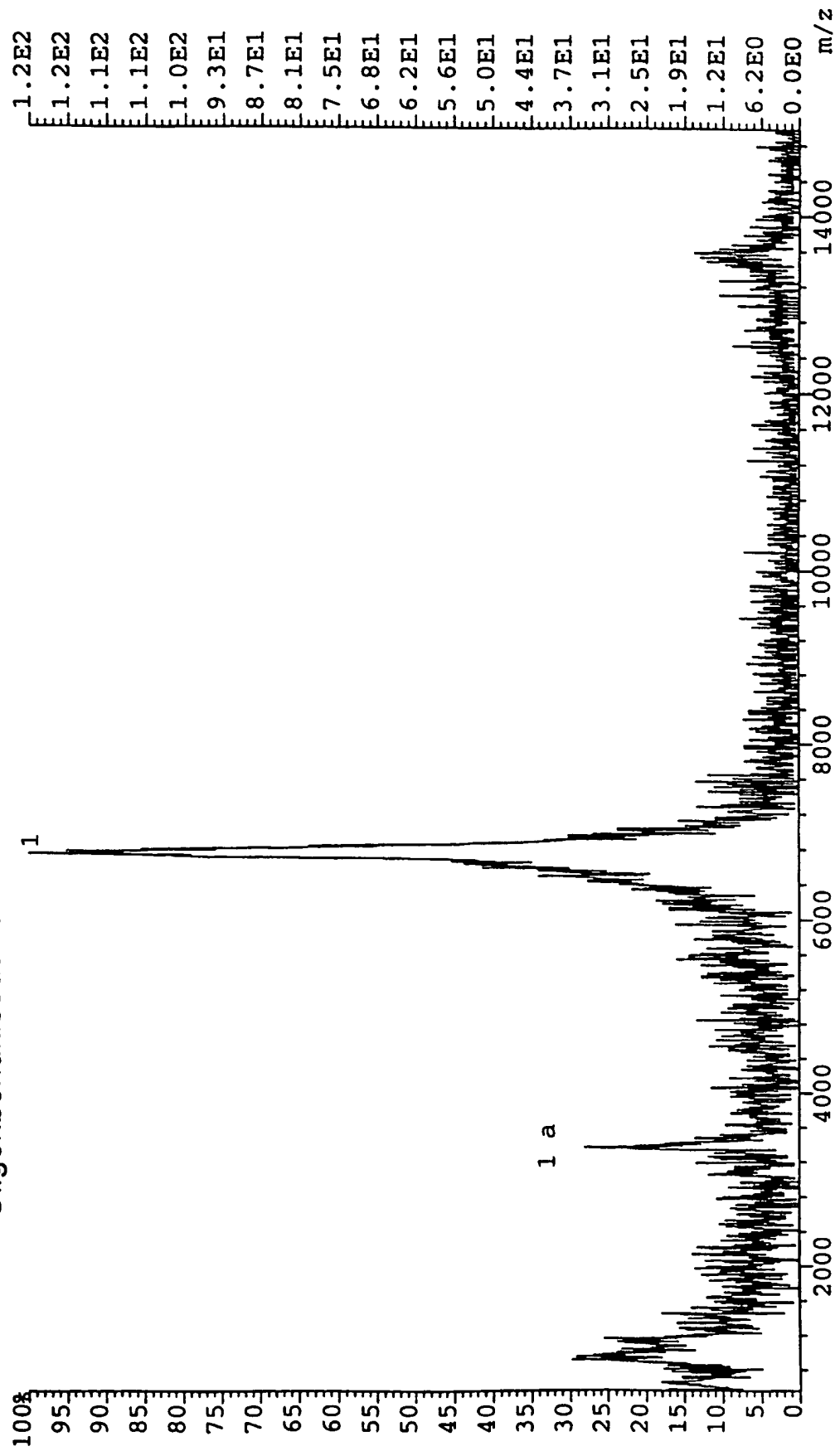
6/17

Figur 4b: Sequenzierung eines 16mers mit 3' -> 5' Phosphodiesterase  
60 Minuten Inkubation



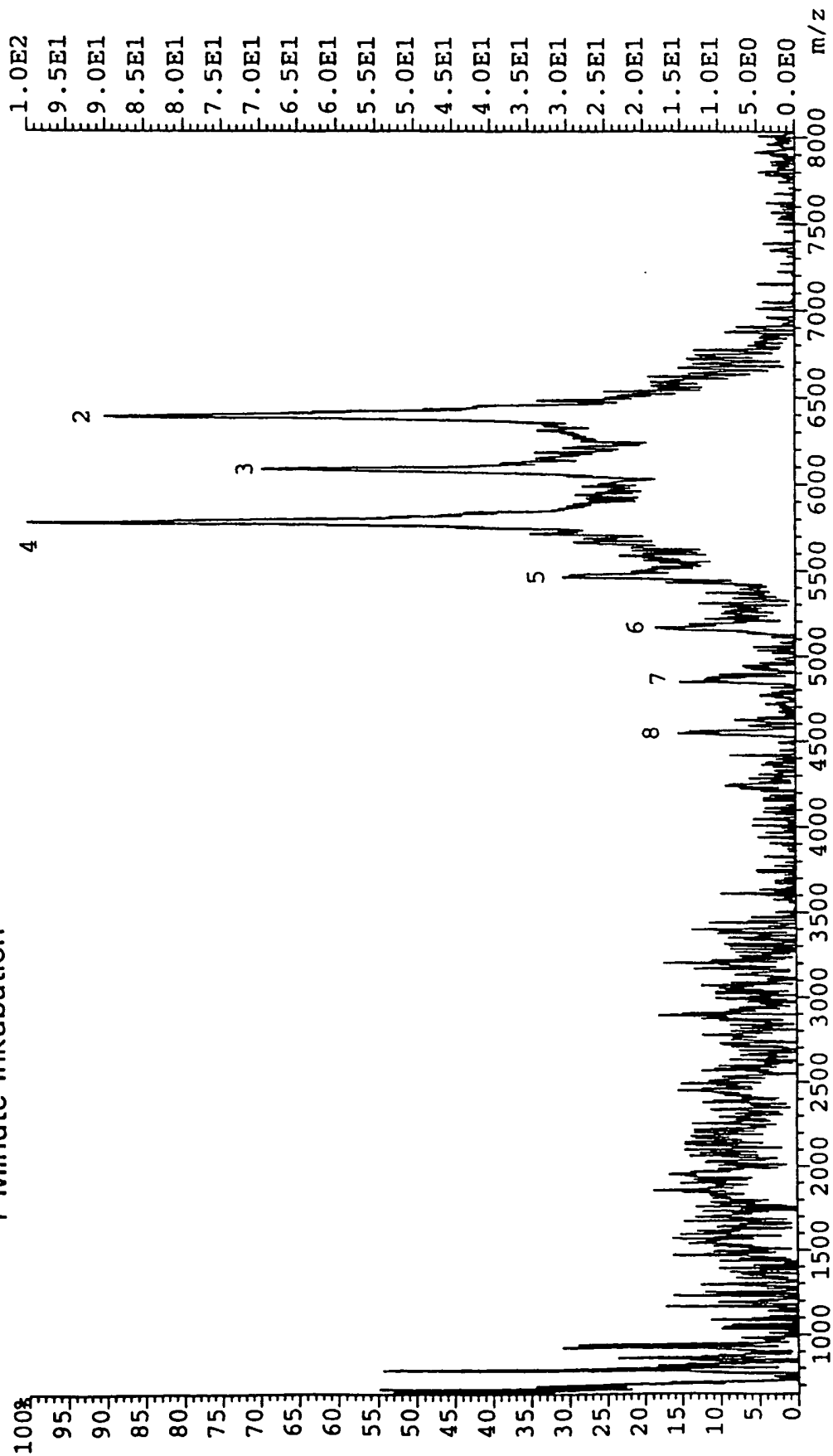
7/17

Figur 5a: Sequenzierung eines 22mers mit 5' -> 3' Phosphodiesterase  
Oligoribonukleotid 2



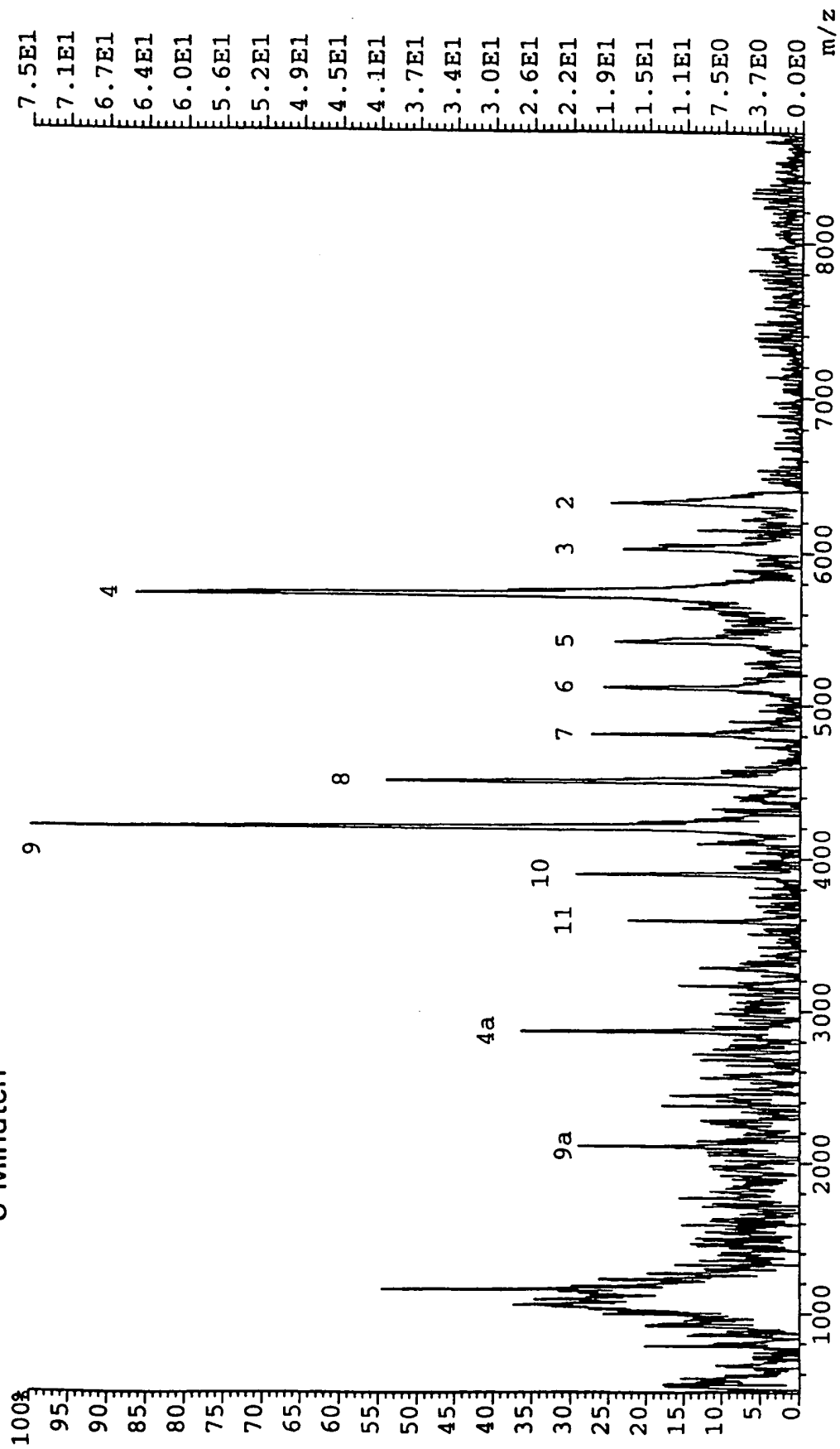
8/17

Figur 5b: Sequenzierung eines 22mers mit 5'->3' Phosphodiesterase  
1 Minute Inkubation



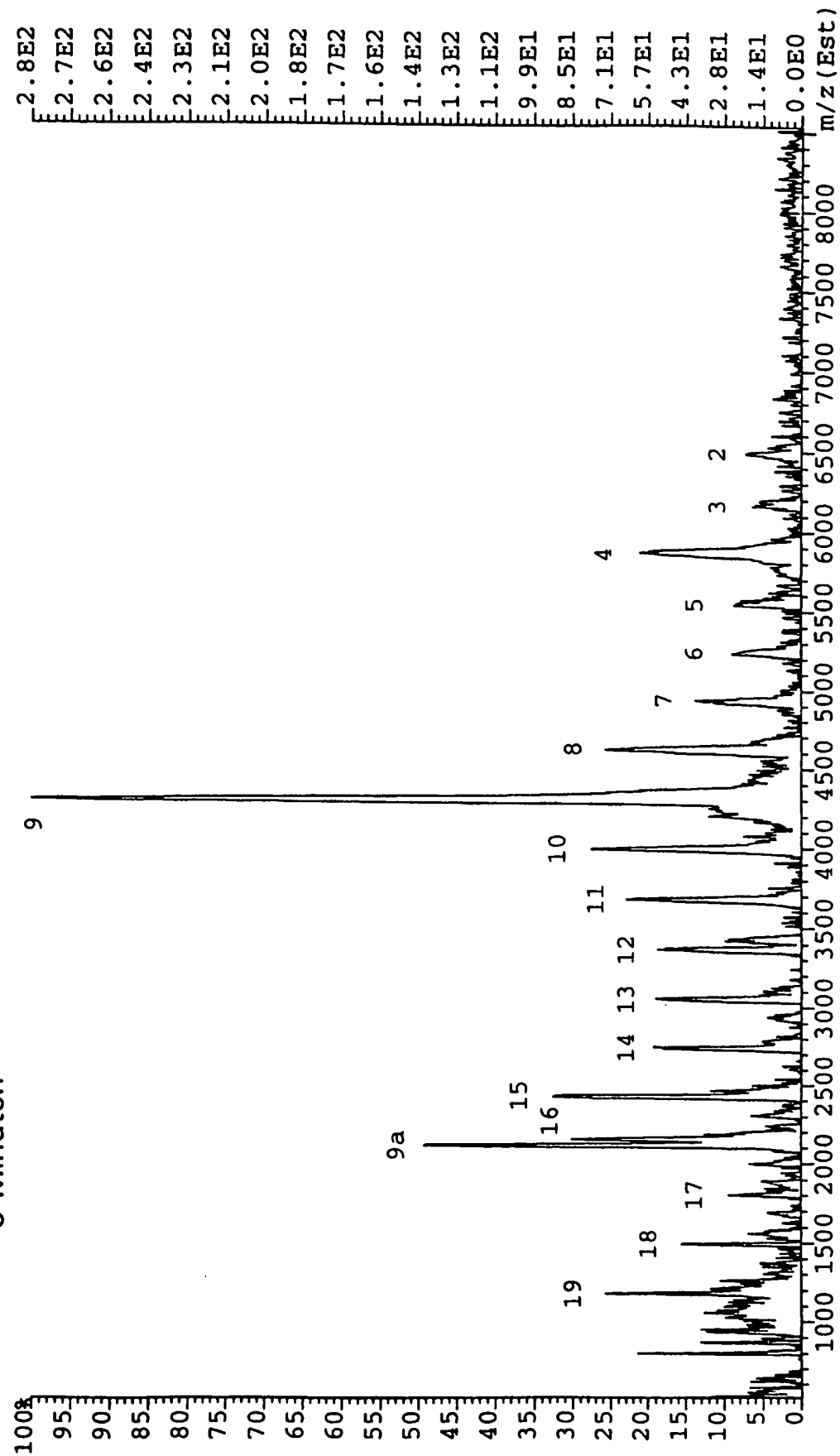
9/17

Figur 5c: Sequenzierung eines 22mers mit 5' -> 3' Phosphodiesterase  
3 Minuten



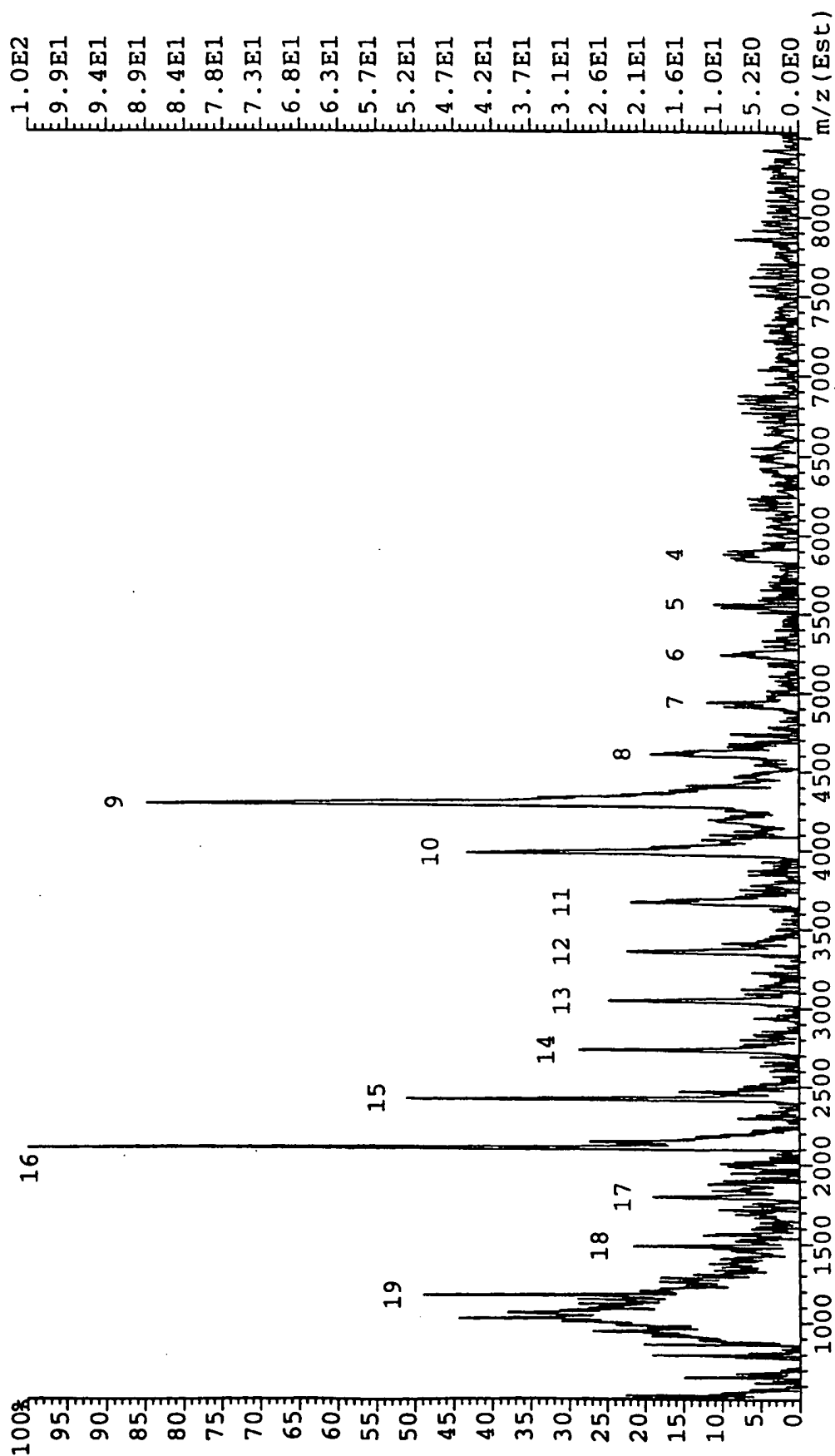
10/17

Figur 5d: Sequenzierung eines 22mers mit 5'->3' Phosphodiesterase  
6 Minuten



11/17

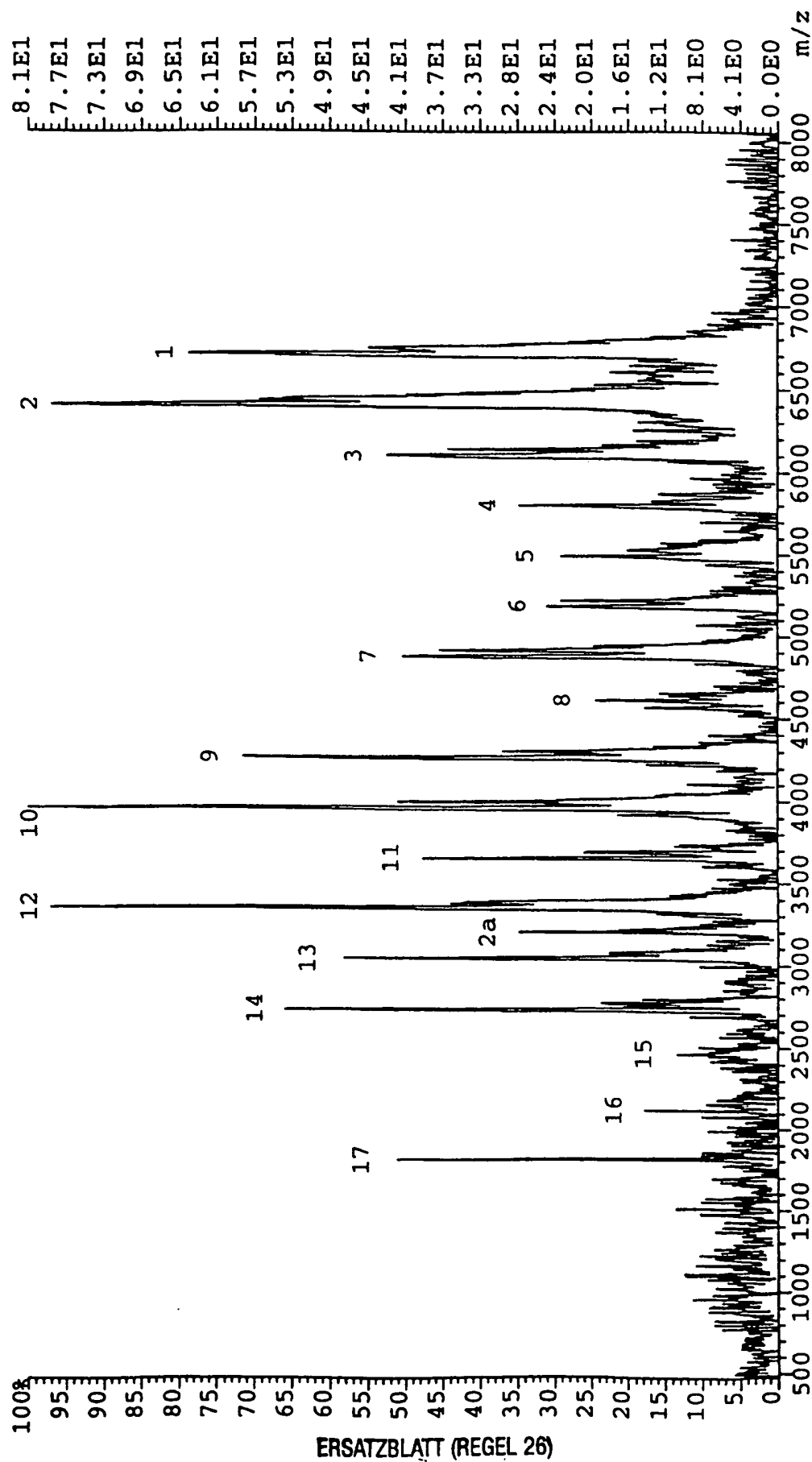
Figur 5e: Sequenzierung eines 22mers mit 5'->3' Phosphodiesterase  
10 Minuten





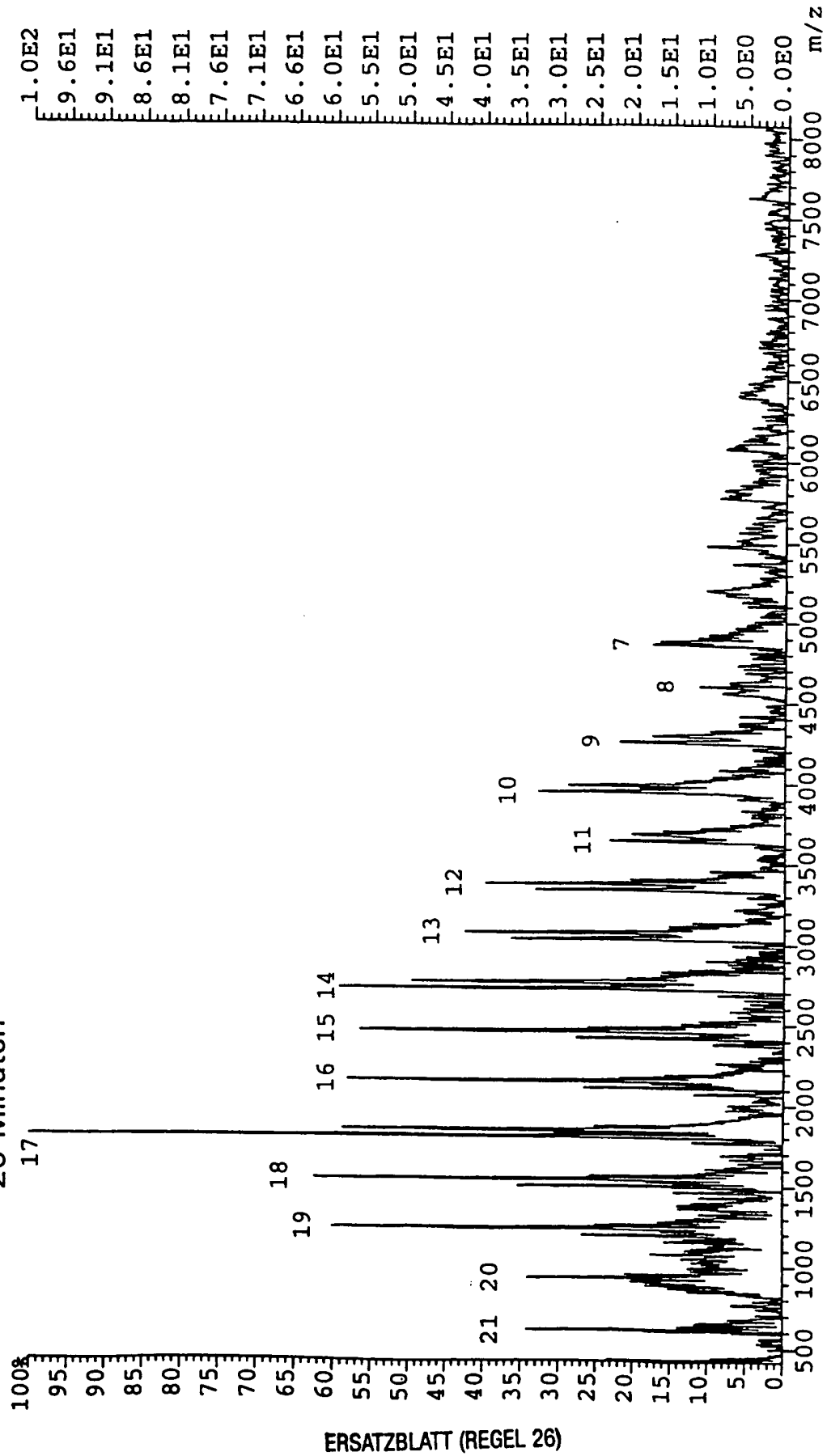
12/17

Figur 6a: Sequenzierung eines 22mers mit 3' -> 5' Phosphodiesterase  
6 Minuten Inkubation

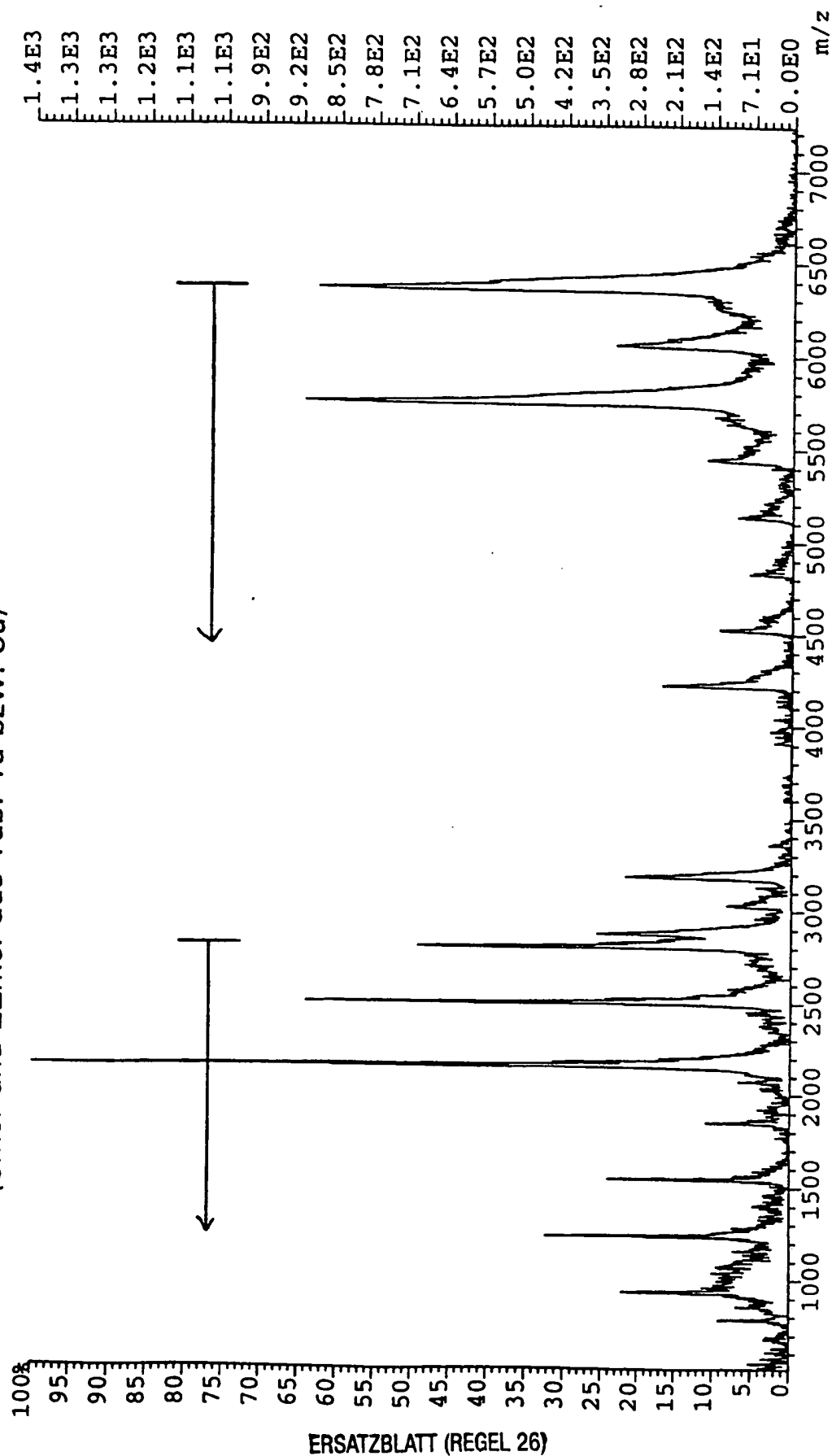


13/17

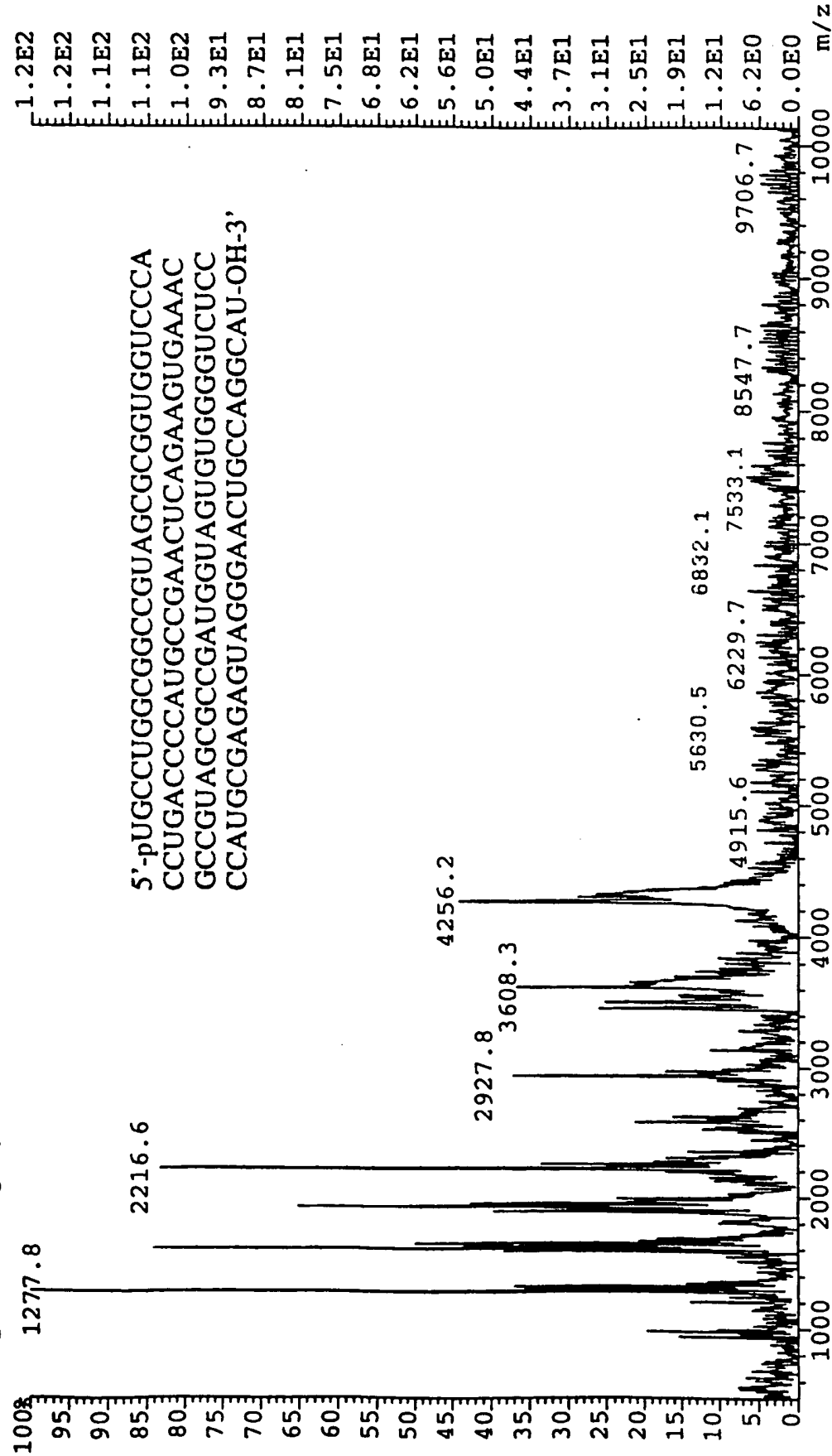
Figur 6b: Sequenzierung eines 22mers mit 3'->5' Phosphodiesterase  
20 Minuten



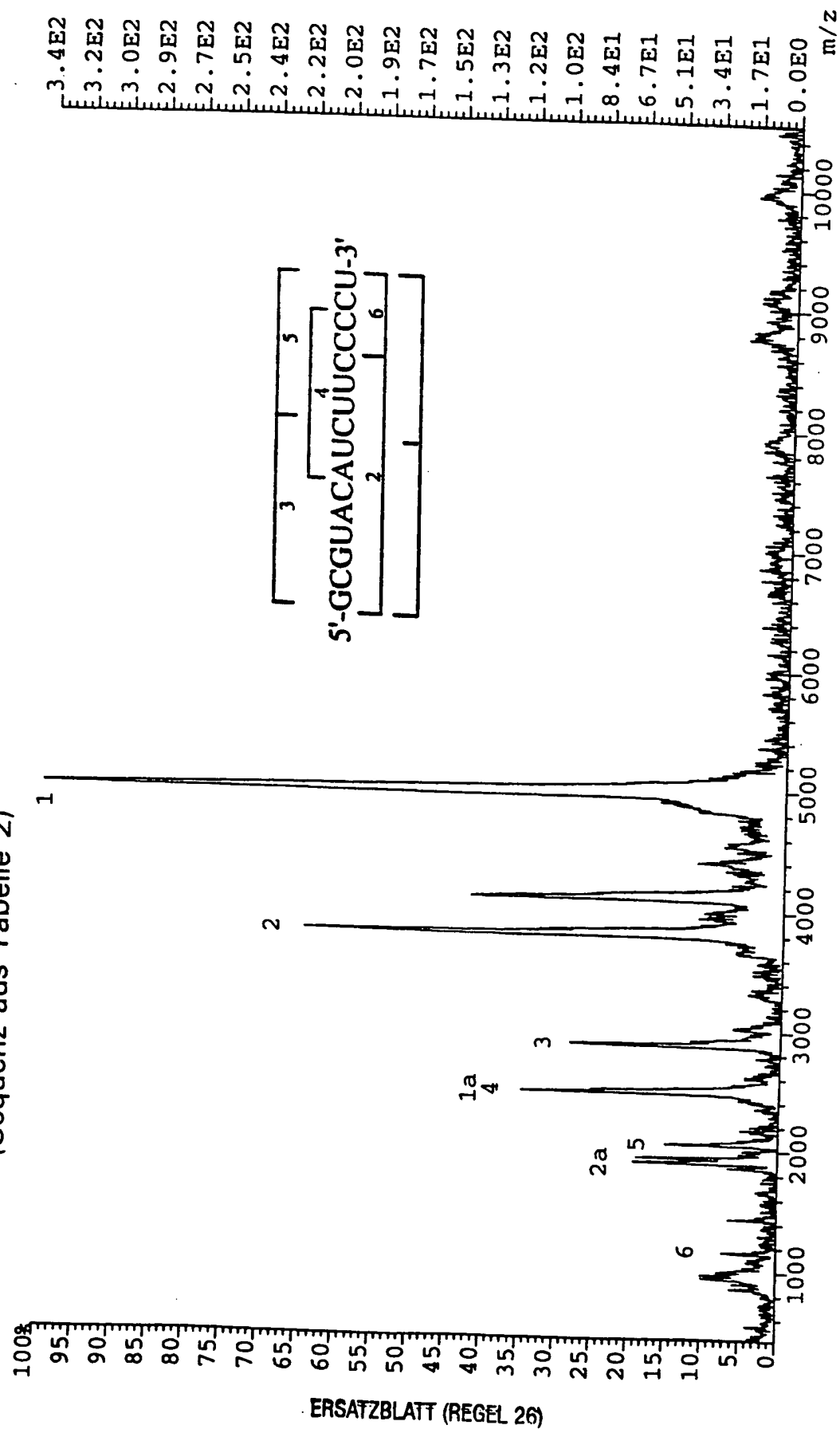
Figur 7: Simultane Sequenzierung von zwei Oligoribonukleotiden  
(8mer und 22mer aus Tab. 1a bzw. 3a)



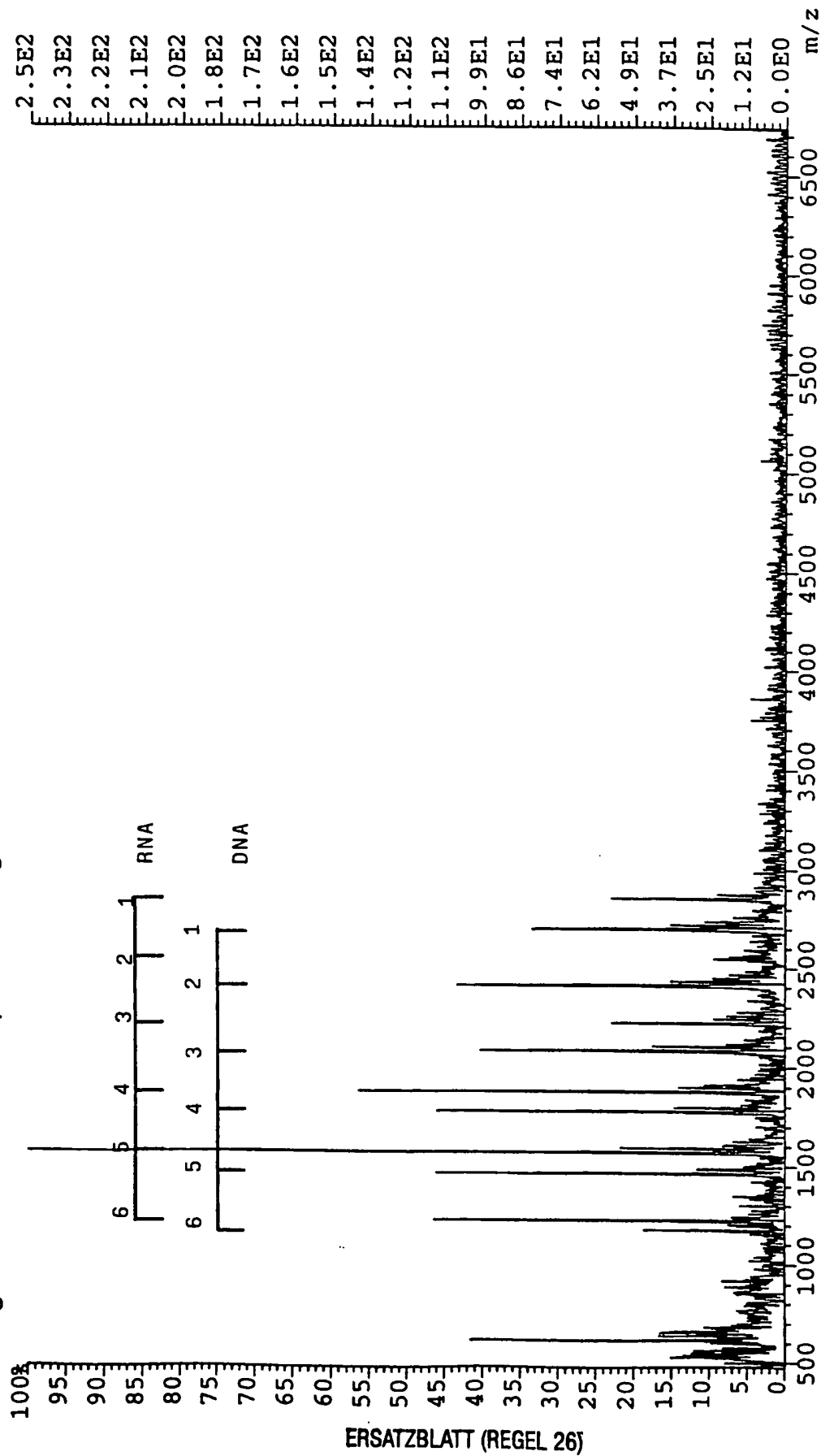
Figur 8: Fingerprint einer 5s-ribosomalen RNA (120mer) mit Rnase CL3



Figur 9: Sequenzierung eines 16mer Oligoribonukleotids (Fingerprint) mit RNase CL3  
(Sequenz aus Tabelle 2)



Figur 10: Simultansequenzierung zweier 9mere (DNA und RNA)



**PCT**WELTORGANISATION FÜR GEISTIGES EIGENTUM  
Internationales BüroINTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

|  |  |   |
|--|--|---|
| <b>(51) Internationale Patentklassifikation <sup>6</sup> :</b><br><b>C12Q 1/68, G01N 33/68</b>   | <b>A3</b>  | <b>(11) Internationale Veröffentlichungsnummer:</b> <b>WO 98/45700</b><br><b>(43) Internationales Veröffentlichungsdatum:</b> 15. Oktober 1998 (15.10.98) |
| <b>(21) Internationales Aktenzeichen:</b> PCT/DE98/01016<br><b>(22) Internationales Anmeldedatum:</b> 8. April 1998 (08.04.98)<br><br><b>(30) Prioritätsdaten:</b><br>197 14 558.2      9. April 1997 (09.04.97)      DE<br><br><b>(71)(72) Anmelder und Erfinder:</b> ENGELS, Joachim, W. [DE/DE]; Feldbergstrasse 1, D-61476 Kronberg (DE). WÖRNER, Karlheinz [DE/DE]; Feldbergstrasse 17, D-63303 Dreieich (DE). FAULSTICH, Konrad [DE/DE]; Brunnenweg 3, D-36115 Ehrenberg (DE). BRILL, Hannelore [DE/DE]; Kastanienweg 7 F, D-61462 Königstein (DE).<br><br><b>(74) Anwalt:</b> SCHULER, Peter; Arnulfstrasse 25, D-80335 München (DE).   | <b>(81) Bestimmungsstaaten:</b> US, europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><br><b>Veröffentlicht</b><br><i>Mit internationalem Recherchenbericht.</i><br><br><b>(88) Veröffentlichungsdatum des internationalen Recherchenrichts:</b> 11. März 1999 (11.03.99) |   |
| <b>(54) Title:</b> METHOD FOR THE MASS SPECTROMETRIC SEQUENCING OF BIOPOLYMERS<br><b>(54) Bezeichnung:</b> VERFAHREN ZUR MASSENSPEKTROMETRISCHEN SEQUENZIERUNG VON BIOPOLYMEREN<br><br><b>(57) Abstract</b><br><p>The invention relates to a new method for sequencing biopolymers by mass spectrometry. The sequencing of biopolymers is either lengthy or requires exact determination of the mass of the fragments, which is difficult, especially in the case of long polymers. The speed of hydrolysis of phosphodiester, -peptide or -glycoside bonds with exo/endonucleases, -peptidases, -glycosidases or other hydrolytically acting substances is used in the inventive method for the sequence analysis of nucleic acids or other biopolymers. Separation and detection of the fragments produced take place by mass spectrometry by determining the mass and the different peak intensities. The main advantage of the method is that an exact determination of the mass is no longer necessary, and also that low-resolution mass spectrometers can be used. Furthermore, analysis and separation of fragments is extremely rapid, as there is no electrophoresis and the sequence of modified nucleic acids can be determined. These principles can also be applied to sequencing or secondary structure determination of other biopolymers, like, for example, peptides and oligosaccharides.</p> |  |   |

#### (57) Zusammenfassung

Es wird ein neues Verfahren zur Sequenzierung von Biopolymeren mit Massenspektrometrie beschrieben. Die Sequenzierung von Biopolymeren ist entweder langwierig oder benötigt eine genaue Massenbestimmung von Fragmenten. Dies ist vor allem für lange Polymere schwierig. Die Geschwindigkeit der Hydrolyse von Phosphodiester-, Peptid- oder Glycosidbindungen mit Exo-/Endonukleasen, -peptidasen, -glycosidasen oder anderen hydrolytisch wirkenden Substanzen wird in unserem Verfahren zur Sequenzanalyse von Nukleinsäuren oder anderen Biopolymeren herangezogen. Trennung und Detektion der erzeugten Fragmente erfolgt mit Massenspektrometrie durch Bestimmung der Masse und unterschiedlichen Peakintensitäten. Der primäre Vorteil der Methode liegt darin, daß keine exakte Massenbestimmung mehr notwendig ist und auch Massenspektrometer mit geringer Auflösung verwendet werden können, Analyse und Trennung der Fragmente extrem schnell sind, da sie elektrophoresefrei sind und die Sequenz modifizierter Nukleinsäuren bestimmt werden kann. Diese Prinzipien lassen sich auch auf die Sequenzierung bzw. Sekundärstrukturbestimmung anderer Biopolymere anwenden, wie z.B. Peptide und Oligosaccharide.

#### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

|    |                              |    |                                   |    |   |    |                                |
|----|------------------------------|----|-----------------------------------|----|---|----|--------------------------------|
| AL | Albanien                     | ES | Spanien                           | LS | Lesotho   | SI | Slowenien                      |
| AM | Armenien                     | FI | Finnland                          | LT | Litauen   | SK | Slowakei                       |
| AT | Österreich                   | FR | Frankreich                        | LU | Luxemburg                                       | SN | Senegal                        |
| AU | Australien                   | GA | Gabun                             | LV | Lettland  | SZ | Swasiland                      |
| AZ | Aserbaidshan                 | GB | Vereinigtes Königreich            | MC | Monaco  | TD | Tschad                         |
| BA | Bosnien-Herzegowina          | GE | Georgien                          | MD | Republik Moldau                                 | TG | Togo                           |
| BB | Barbados                     | GH | Ghana                             | MG | Madagaskar                                      | TJ | Tadschikistan                  |
| BE | Belgien                      | GN | Guinea                            | MK | Die ehemalige jugoslawische Republik Mazedonien | TM | Turkmenistan                   |
| BF | Burkina Faso                 | GR | Griechenland                      | ML | Mali  | TR | Türkei                         |
| BG | Bulgarien                    | HU | Ungarn                            | MN | Mongolei  | TT | Trinidad und Tobago            |
| BJ | Benin                        | IE | Irland                            | MR | Mauretanien                                     | UA | Ukraine                        |
| BR | Brasilien                    | IL | Israel                            | MW | Malawi  | UG | Uganda                         |
| BY | Belarus                      | IS | Island                            | MX | Mexiko  | US | Vereinigte Staaten von Amerika |
| CA | Kanada                       | IT | Italien                           | NE | Niger   | UZ | Usbekistan                     |
| CF | Zentralafrikanische Republik | JP | Japan                             | NL | Niederlande                                     | VN | Vietnam                        |
| CG | Kongo                        | KE | Kenia                             | NO | Norwegen  | YU | Jugoslawien                    |
| CH | Schweiz                      | KG | Kirgisistan                       | NZ | Neuseeland                                      | ZW | Zimbabwe                       |
| CI | Côte d'Ivoire                | KP | Demokratische Volksrepublik Korea | PL | Polen   |    |                                |
| CM | Kamerun                      | KR | Republik Korea                    | PT | Portugal  |    |                                |
| CN | China                        | KZ | Kasachstan                        | RO | Rumänien  |    |                                |
| CU | Kuba                         | LC | St. Lucia                         | RU | Russische Föderation                            |    |                                |
| CZ | Tschechische Republik        | LI | Liechtenstein                     | SD | Sudan   |    |                                |
| DE | Deutschland                  | LK | Sri Lanka                         | SE | Schweden  |    |                                |
| DK | Dänemark                     | LR | Liberia                           | SG | Singapur  |    |                                |
| EE | Estland                      |    |                                   |    |   |    |                                |



# INTERNATIONAL SEARCH REPORT

International Application No

PCT/DE 98/01016

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6 C12Q1/68 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|------------|---|-----------------------|
| E          | WO 98 20166 A (DEN BOOM DIRK VAN ; JURINKE CHRISTIAN (DE); HIGGINS G SCOTT (DE); L)<br>14 May 1998<br>see the whole document<br>---   | 1-8                   |
| P, X       | WO 98 03684 A (HYBRIDON INC)<br>29 January 1998<br>see page 4, line 2 - line 23<br>---  | 1-8                   |
| P, X       | WO 97 33000 A (GENETRACE SYSTEMS INC)<br>12 September 1997<br>see claims<br>see page 4, line 26 - page 5, line 18<br>see page 11, line 21 - page 12, line 18<br>see page 17, line 25 - page 18, line 5<br>---<br>-/-- | 1-8                   |

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

26 October 1998

Date of mailing of the international search report

01/12/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Routledge, 8

## INTERNATIONAL SEARCH REPORT

Inter:      nal Application No

PCT/DE 98/01016

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document with indication, where appropriate, of the relevant passages  | Relevant to claim No |
|----------|--|----------------------|
| X        | WO 96 36986 A (PERSEPTIVE BIOSYSTEMS INC)<br>21 November 1996<br>see claims<br>see page 3, line 18 - line 28<br>see page 4, line 18 - line 25<br>see page 13, line 1 - line 4<br>see page 17, line 24 - page 20, line 22<br>see example 1<br>--- | 1-8                  |
| X        | WO 94 21822 A (KOESTER HUBERT)<br>29 September 1994<br>see the whole document<br>---   | 1-8                  |
| X        | WO 90 04596 A (3I RESEARCH EXPLOIT LTD)<br>3 May 1990<br>see claims<br>see page 2, paragraph 3<br>see page 10, paragraph 2 - paragraph 3<br>---  | 1-8                  |
| X        | WOERNER K ET AL: "RECENT DEVELOPMENTS IN<br>MASS SPECTROMETRY: RNA SEQUENCING"<br>NUCLEOSIDES & NUCLEOTIDES,<br>vol. 16, no. 5/06, 1997, pages 573-577,<br>XP000198057<br>see the whole document<br>-----  | 1-8                  |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/DE 98/01016

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---------------------|----------------------------|---------------------|
| WO 9820166 A                              | 14-05-1998          | AU 5106998 A               | 29-05-1998          |
|   |                     | AU 5247298 A               | 29-05-1998          |
|   |                     | WO 9820019 A               | 14-05-1998          |
|   |                     | AU 5198098 A               | 29-05-1998          |
|   |                     | WO 9820020 A               | 14-05-1998          |
| WO 9803684 A                              | 29-01-1998          | AU 4042597 A               | 10-02-1998          |
| WO 9733000 A                              | 12-09-1997          | AU 2069597 A               | 22-09-1997          |
| WO 9636986 A                              | 21-11-1996          | EP 0827628 A               | 11-03-1998          |
| WO 9421822 A                              | 29-09-1994          | AU 687801 B                | 05-03-1998          |
|   |                     | AU 6411694 A               | 11-10-1994          |
|   |                     | CA 2158642 A               | 29-09-1994          |
|   |                     | EP 0689610 A               | 03-01-1996          |
|   |                     | JP 8507926 T               | 27-08-1996          |
|   |                     | US 5622824 A               | 22-04-1997          |
| WO 9004596 A                              | 03-05-1990          | EP 0440732 A               | 14-08-1991          |

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
IPK 6 C12Q1/68 G01N33/68

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
IPK 6 C12Q G01N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

| Kategorie* | Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile   | Betr. Anspruch Nr. |
|------------|--|--------------------|
| E          | WO 98 20166 A (DEN BOOM DIRK VAN ; JURINKE CHRISTIAN (DE); HIGGINS G SCOTT (DE); L)<br>14. Mai 1998<br>siehe das ganze Dokument<br>---   | 1-8                |
| P, X       | WO 98 03684 A (HYBRIDON INC)<br>29. Januar 1998<br>siehe Seite 4, Zeile 2 - Zeile 23<br>---  | 1-8                |
| P, X       | WO 97 33000 A (GENETRACE SYSTEMS INC)<br>12. September 1997<br>siehe Ansprüche<br>siehe Seite 4, Zeile 26 - Seite 5, Zeile 18<br>siehe Seite 11, Zeile 21 - Seite 12, Zeile 18<br>siehe Seite 17, Zeile 25 - Seite 18, Zeile 5<br>---<br>-/- | 1-8                |

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

\* Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"Z" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

26. Oktober 1998

Absendedatum des internationalen Recherchenberichts

01/12/1998

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Routledge, B

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

| Kategorie | Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile  | Betr. Anspruch Nr. |
|-----------|---|--------------------|
| X         | WO 96 36986 A (PERSEPTIVE BIOSYSTEMS INC)<br>21. November 1996<br>siehe Ansprüche<br>siehe Seite 3, Zeile 18 - Zeile 28<br>siehe Seite 4, Zeile 18 - Zeile 25<br>siehe Seite 13, Zeile 1 - Zeile 4<br>siehe Seite 17, Zeile 24 - Seite 20, Zeile 22<br>siehe Beispiel 1<br>---- | 1-8                |
| X         | WO 94 21822 A (KOESTER HUBERT)<br>29. September 1994<br>siehe das ganze Dokument<br>----  | 1-8                |
| X         | WO 90 04596 A (3I RESEARCH EXPLOIT LTD)<br>3. Mai 1990<br>siehe Ansprüche<br>siehe Seite 2, Absatz 3<br>siehe Seite 10, Absatz 2 - Absatz 3<br>----   | 1-8                |
| X         | WOERNER K ET AL: "RECENT DEVELOPMENTS IN<br>MASS SPECTROMETRY: RNA SEQUENCING"<br>NUCLEOSIDES & NUCLEOTIDES,<br>Bd. 16, Nr. 5/06, 1997, Seiten 573-577,<br>XP000198057<br>siehe das ganze Dokument<br>-----   | 1-8                |

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/DE 98/01016

| Im Recherchenbericht<br>angeführtes Patentdokument | Datum der<br>Veröffentlichung | Mitglied(er) der<br>Patentfamilie | Datum der<br>Veröffentlichung |
|--|-------------------------------|-----------------------------------|-------------------------------|
| WO 9820166 A                                       | 14-05-1998                    | AU 5106998 A                      | 29-05-1998                    |
|  |                               | AU 5247298 A                      | 29-05-1998                    |
|  |                               | WO 9820019 A                      | 14-05-1998                    |
|  |                               | AU 5198098 A                      | 29-05-1998                    |
|  |                               | WO 9820020 A                      | 14-05-1998                    |
| WO 9803684 A                                       | 29-01-1998                    | AU 4042597 A                      | 10-02-1998                    |
| WO 9733000 A                                       | 12-09-1997                    | AU 2069597 A                      | 22-09-1997                    |
| WO 9636986 A                                       | 21-11-1996                    | EP 0827628 A                      | 11-03-1998                    |
| WO 9421822 A                                       | 29-09-1994                    | AU 687801 B                       | 05-03-1998                    |
|  |                               | AU 6411694 A                      | 11-10-1994                    |
|  |                               | CA 2158642 A                      | 29-09-1994                    |
|  |                               | EP 0689610 A                      | 03-01-1996                    |
|  |                               | JP 8507926 T                      | 27-08-1996                    |
|  |                               | US 5622824 A                      | 22-04-1997                    |
| WO 9004596 A                                       | 03-05-1990                    | EP 0440732 A                      | 14-08-1991                    |